CROSS-DISPLAY OBJECT MOVEMENT IN MULTI-DISPLAY ENVIRONMENTS

A Thesis Submitted to the College of
Graduate Studies and Research
In Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy
In the Department of Computer Science
University of Saskatchewan
Saskatoon

By

Miguel Ángel Nacenta Sánchez

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Computer Science
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5C9

PERMISSIONS

Parts of this dissertation have been published in other forms; copyrights of those works are held by the Association of Computer Machinery and Taylor and Francis. In all cases, permission has been granted to include these materials in this dissertation.

ASSOCIATION FOR COMPUTING MACHINERY

Chapter 6

Nacenta, M. A., Aliakseyeu, D., Subramanian, S., and Gutwin, C. 2005. A comparison of techniques for multi-display reaching. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA, April 02 - 07, 2005). CHI '05. ACM, New York, NY, 371-380. DOI= http://doi.acm.org/10.1145/1054972.1055024

Nacenta, M. A., Sallam, S., Champoux, B., Subramanian, S., and Gutwin, C. 2006. Perspective cursor: perspective-based interaction for multi-display environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada, April 22 - 27, 2006). R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, and G. Olson, Eds. CHI '06. ACM, New York, NY, 289-298. DOI= http://doi.acm.org/10.1145/1124772.1124817

Chapter 7

Nacenta, M. A., Mandryk, R. L., and Gutwin, C. 2008. Targeting across displayless space. In *Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy, April 05 - 10, 2008). CHI '08. ACM, New York, NY, 777-786. DOI= http://doi.acm.org/10.1145/1357054.1357178

TAYLOR AND FRANCIS

Taylor and Francis grants the author the right to include an article in a thesis or dissertation that is not to be published commercially, provided that acknowledgement to prior publication in the relevant Taylor & Francis journal is made explicit. Parts of the following article have been used in Chapters 1 through 7.

Nacenta, M. A., Gutwin, C., Aliakseyeu, D. & Subramanian, S. (2009). There and Back Again: Cross-Display Object Movement in Multi-Display Environments. *Human-Computer Interaction*, *24*(1), 170-229. doi:10.1080/07370020902819882

FIGURES

Figures 23 and 27 reproduced with permission from ACM from [Johanson et al. 2002], [Shoemaker et al., 2007] respectively.

Figures 24, 25, and 28 reproduced with permission from the authors from [Biehl and Bailey, 2004], [Baudisch et al, 2003] and [Hinckley et al., 2006] respectively.

ABSTRACT

Many types of multi-display environments (MDEs) are emerging that allow users to better interact with computers. In these environments, being able to move visual objects (such as window icons or the cursor) from one display to another is a fundamental activity.

This dissertation focuses on understanding how human performance of cross-display actions is affected by the design of cross-display object movement interaction techniques. Three main aspects of cross-display actions are studied: how displays are referred to by the system and the users, how spatial actions are planned, and how actions are executed. Each of these three aspects is analyzed through laboratory experiments that provide empirical evidence on how different characteristics of interaction techniques affect performance.

The results further our understanding of cross-display interaction and can be used by designers of new MDEs to create more efficient multi-display interfaces.

ACKNOWLEDGMENTS

I would like to express my gratitude for the guidance, patience and unconditional support received from Dr. Carl Gutwin throughout this learning process. It is fortunate for a Ph.D. student to be supervised by a prominent figure in the field such as Carl, but it is crucial to receive guidance beyond the strictly academic. I cannot exaggerate the quality of the support received from Carl in the last five years in the scientific, academic, ethical and personal spheres. It is going to be difficult to give back in the amounts received; however, if I can pass forward a fraction of what I have learnt, I am confident I can make a worthy contribution.

I must especially thank Sriram Subramanian, who was my co-supervisor during the initial years of the Ph.D.; Sriram brought me to Saskatchewan and has been a constant source of exceptional advice and encouragement before and throughout this process. I was lucky to meet Sriram and to have my life affected so fundamentally by him.

Much credit is due as well to the members of my committee: Nadeem Jamali, Kevin Schneider, Regan Mandryk, Lorin Elias, and Brian Bailey. The local members of the committee have provided useful feedback that has helped to improve the document, especially at the proposal stage. Regan deserves a special mention since she has been a fantastic collaborator and a model to follow.

Other collaborators have been crucial to the work contained in this dissertation, but also in other publications that have been informative for my dissertation, but are not directly included in this document: Dima Aliakseyeu, Mutasem Barjawi, Scott Bateman, Patrick Baudisch, Hrvoje Benko, Bernard Champoux, Yuichi Itoh, Fumio Kishino, Yoshifumi Kitamura, Yohei Miki, David Pinelle, Adrian Reetz, Satoshi Sakurai, Samer Sallam, Tad Stach, Dane Stuckel, Tokuo Yamaguchi and Andy Wilson.

On a personal note, I am grateful for my parents' continuous support and for planting in me the seed of curiosity and academic endeavor (¡gracias por todo!) and, of course, to Kate for making this all meaningful (as well as the grammar). Finally, I would like to thank Gwen, for letting me steal Carl for too many nights; Jan Thompson, for all the support; Daniela and Lindell, for all the fun during the Saskatoon winters; and Sheelagh Carpendale, for support and advice this last year.

Dedication

Para Ángeles, Miguel Ángel, y Kate

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF VIDEO FIGURES

PUBLICATIONS OF THE AUTHOR

Publications with content from this dissertation

Nacenta, Miguel A., Dzmitry Aliakseyeu, Sriram Subramanian, and Carl Gutwin. 2005. A comparison of techniques for multi-display reaching. In Proceedings of the SIGCHI conference on Human factors in computing systems, 371-380. Portland, Oregon, USA: ACM. doi:10.1145/1054972.1055024. http://portal.acm.org/citation.cfm?doid=1055024.

Nacenta, Miguel A., Samer Sallam, Bernard Champoux, Sriram Subramanian, and Carl Gutwin. 2006. Perspective cursor: perspective-based interaction for multi-display environments. In Proceedings of the SIGCHI conference on Human Factors in computing systems, 289-298. Montréal, Québec, Canada: ACM. doi:10.1145/1124772.1124817. http://portal.acm.org/citation.cfm?id=1124772.1124817.

Nacenta, Miguel A., Regan L. Mandryk, and Carl Gutwin. 2008. Targeting across displayless space. In Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, 777-786. Florence, Italy: ACM. doi:10.1145/1357054.1357178. http://portal.acm.org/citation.cfm?id=1357178.

Nacenta, Miguel A., C. Gutwin, D. Aliakseyeu, and S. Subramanian. 2009. There and Back Again: Cross-Display Object Movement in Multi-Display Environments. Human-Computer Interaction 24, no. 1: 170-229.

Publications related to this dissertation

Nacenta, Miguel A., David Pinelle, Dane Stuckel, and Carl Gutwin. 2007. The effects of interaction technique on coordination in tabletop groupware. In Proceedings of Graphics

Interface 2007, 191-198. Montreal, Canada: ACM. doi:10.1145/1268517.1268550. http://portal.acm.org/citation.cfm?id=1268550.

Nacenta, Miguel A., Satoshi Sakurai, Tokuo Yamaguchi, Yohei Miki, Yuichi Itoh, Yoshifumi Kitamura, Sriram Subramanian, and Carl Gutwin. 2007. E-conic: a perspective-aware interface for multi-display environments. In Proceedings of the 20th annual ACM symposium on User interface software and technology, 279-288. Newport, Rhode Island, USA: ACM. doi:10.1145/1294211.1294260. http://portal.acm.org/citation.cfm?id=1294260.

Aliakseyeu, Dzmitry, **Miguel A. Nacenta**, Sriram Subramanian, and Carl Gutwin. 2006. Bubble radar: efficient pen-based interaction. In Proceedings of the working conference on Advanced visual interfaces, 19-26. Venezia, Italy: ACM. doi:10.1145/1133265.1133271. http://portal.acm.org/citation.cfm?id=1133265.1133271.

Reetz, Adrian, Carl Gutwin, Tadeusz Stach, **Miguel A. Nacenta**, and Sriram Subramanian. 2006. Superflick: a natural and efficient technique for long-distance object placement on digital tables. In Proceedings of Graphics Interface 2006, 163-170. Quebec, Canada: Canadian Information Processing Society. http://portal.acm.org/citation.cfm?id=1143079.1143106.

Pinelle, David, **Miguel A. Nacenta**, Carl Gutwin, and Tadeusz Stach. 2008. The effects of co-present embodiments on awareness and collaboration in tabletop groupware. In Proceedings of graphics interface 2008, 1-8. Windsor, Ontario, Canada: Canadian Information Processing Society. http://portal.acm.org/citation.cfm?id=1375716.

Sakurai, S., Y. Itoh, Y. Kitamura, **Miguel A. Nacenta**, T. Yamaguchi, S. Subramanian, and F. Kishino. 2008. A Middleware for Seamless Use of Multiple Displays. Lecture Notes In Computer Science 5136: 252-266.

Sakurai, Satoshi, Tokuo Yamaguchi, Yoshifumi Kitamura, Yuichi Itoh, Ryo Fukazawa, Fumio Kishino, **Miguel A. Nacenta**, and Sriram Subramanian. 2008. M3: multi-modal interface in multi-display environment for multi-users. In ACM SIGGRAPH ASIA 2008 artgallery:

emerging technologies, 45-45. Singapore: ACM. doi:10.1145/1504229.1504259. http://portal.acm.org/citation.cfm?id=1504259.

Pinelle, David, Mutasem Barjawi, **Miguel A. Nacenta**, and Regan Mandryk. 2009. An evaluation of coordination techniques for protecting objects and territories in tabletop groupware. In Proceedings of the 27th international conference on Human factors in computing systems, 2129-2138. Boston, MA, USA: ACM. doi:10.1145/1518701.1519025. http://portal.acm.org/citation.cfm?id=1519025.

Miguel A. Nacenta, David Pinelle, Carl Gutwin, Regan Mandryk, 2010. Individual and Grooup Support in Tabletop Interaction Techniques. In Müller-Tomfelde, C. (ed.) Tabletops - Horizontal Interactive Displays. Human Computer Interaction Series, Springer Verlag, 2010.

Other Publications

Bateman, Scott, Carl Gutwin, and **Miguel A. Nacenta**. 2008. Seeing things in the clouds: the effect of visual features on tag cloud selections. In Proceedings of the nineteenth ACM conference on Hypertext and hypermedia, 193-202. Pittsburgh, PA, USA: ACM. doi:10.1145/1379092.1379130. http://portal.acm.org/citation.cfm?id=1379092.1379130.

Marquardt, Nicolai, **Miguel A. Nacenta**, James E. Young, Sheelagh Carpendale, Saul Greenberg, Ehud Sharlin, 2009. The Haptic Tabletop Puck: Tactile Feedback for Interactive Tabletops. In Proceedings of the first international conference on Interactive Tabletops and Surfaces (ACM ITS - Tabletop 2009), 93-100. ACM Press.

Hancock, Mark, **Miguel A. Nacenta**, Carl Gutwin, Sheelagh Carpendale, 2009. The Effects of Changing Projection Geometry on the Interpretation of 3D Orientation on Tabletops. In Proceedings of the first international conference on Interactive Tabletops and Surfaces (ACM ITS - Tabletop 2009), 175-182. ACM Press.

CHAPTER 1: INTRODUCTION

Multi-display environments (MDEs) are becoming more and more common, and are moving beyond simple multi-monitor setups to more complex environments that link tabletops, wall displays, projectors, PC monitors, and mobile devices into a single workspace. These large-scale MDEs have the potential to dramatically change the way that we work with digital information: for example, they provide a variety of work surfaces to fit different kinds of tasks, they provide a very large display surface, they enable the use of peripheral attention space, and they naturally support co-located collaboration.

Most current multi-display interfaces are superficial adaptations of single-display interfaces; however, MDEs are substantially different from single-display systems and require specific interaction techniques to perform multi-display operations. One of the most important multi-display actions is cross-display object movement – the action of moving the representation of a digital object (e.g., an icon, a window or the cursor) from one display to another. Without interaction techniques that allow object movement between displays there can be little multi-display interaction.

This work is focused on the study of interaction techniques for cross-display object movement[1], i.e., on the parts of the user interface that a system implements to let users accomplish the movement of visual objects from one display to another. Although moving an object between displays is a simple action, there are many ways in which users can perform it; for example, users could physically touch the destination display, select it from a list, manipulate a mouse to remotely move the object around the entire environment, or press a key several times until the object appears in the right display. As we will see, subtle characteristics of these interaction techniques can have a large impact on the overall efficiency of the system and on the user experience.

[1] For the remainder of this dissertation, unless stated otherwise, the phrase “object” will be taken to mean “visual object”, i.e., the visual representation on screen of a digital object, regardless of where the actual information of the object is stored in memory (the information associated with the object could reside in any device that is part of the MDE).

The goal of this dissertation is to increase the understanding of the factors that affect performance in cross-display object movement (CDOM), and thereby to improve the interaction performance in MDEs. In particular, I focus on a set of three characteristics of interaction techniques that are related to a basic model of the cross-display action (i.e., how to refer to displays, how to plan the control movement, and how to execute the control movement) and analyze how these will affect performance, user preference, and subjective workload. A better understanding of these issues is intended to help researchers and designers to design better interaction techniques and improve the interfaces of increasingly complex MDEs.

1.1. An Example of the Problem of Cross-display Object Movement

The interaction techniques that MDEs implement for CDOM can result in inadequate interface designs. For example, consider the smart office scenario of Figure 1.Left, in which user A intends to move a window to a laptop on another table. By using a technique similar to Multibrowsing [Johanson et al., 2001], user A would have to invoke a contextual cross-display menu by right clicking on the window, and then selecting the right destination display from a list of display names. This mechanism requires that users remember each display by name, which takes added time and effort, especially if there are many displays in the system. Most of the time, users will simply want to move content to a particular display which is easy to identify (for example, by a pointing gesture) instead of choosing an arbitrarily-selected name from a list of labels such as the one represented in Figure 1. Moreover, using this technique, other people in the room will have very little information available about who performed the action or from which display.

The increasing complexity of current MDEs can also make traditional solutions to the cross-display object transfer obsolete. For example, the usual mechanism to bind two displays in a dual-monitor configuration is to combine all physical display space into a single, larger virtual space without boundaries (Figure 2.A), allowing the cursor to move from one display to another by crossing the boundary between them. However, this approach does not scale to advanced multi-display scenarios such as smart offices because there are too many displays of different sizes in different orientations that are not easily represented as a single flat surface (Figure 2.B). Similarly, combining displays in this way becomes impractical for users that need to move around the environment because there are many alternative flat representations of the 3D physical environment and the correct representation depends on the point of view – see Figure 3.

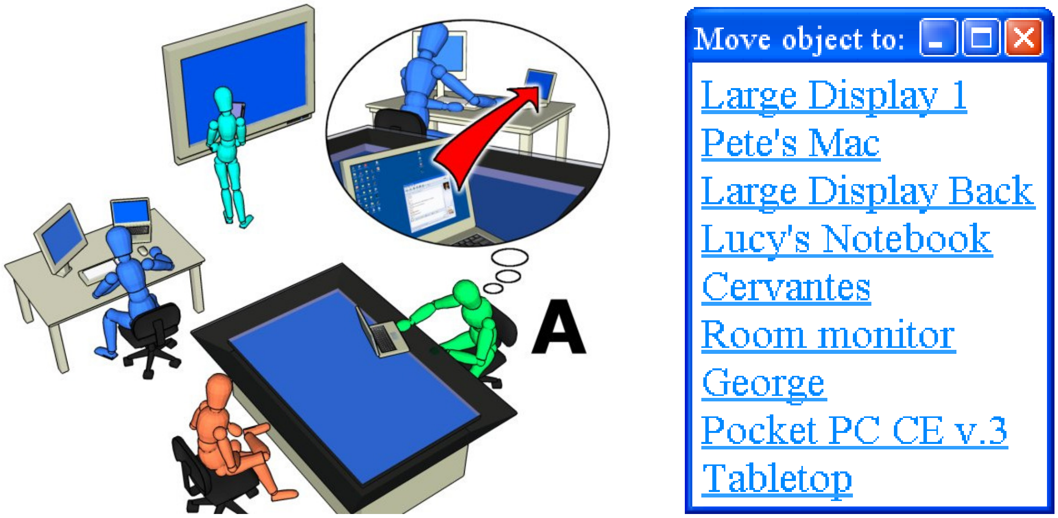

Figure 1. Left) User A wishes to move an object to a nearby display. Right) The user must select the destination display from a list of display labels (similar to [Johanson et al., 2001]).

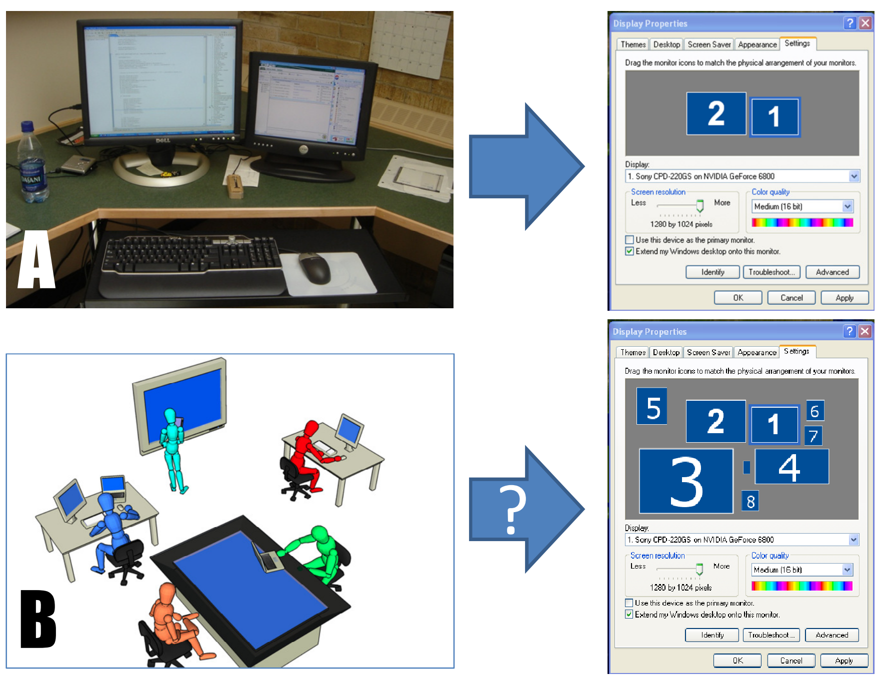

Figure 2. A) dual-monitor display and its virtual space mapping. B) A complex multi-display environment and a possible (but unusable) space mapping.

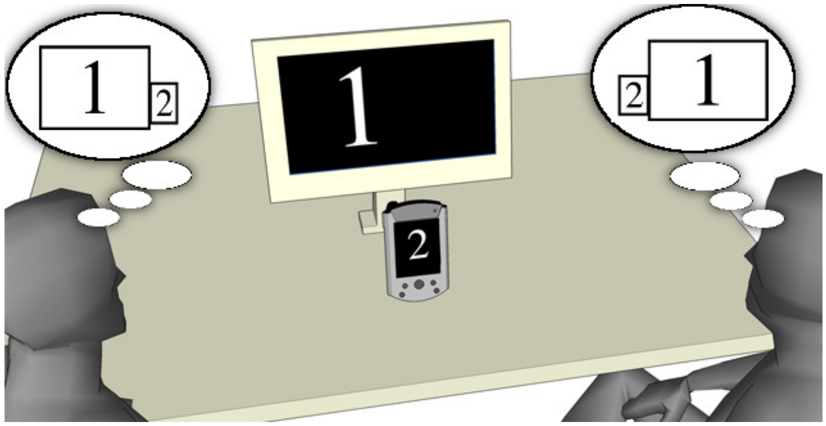

Figure 3. The "correct" planar mapping of an MDE can depend on the position of the user.

1.2. Multi-display Environments

For this dissertation I will define a multi-display environment (also known as a multi-surface environment – MSE) as an interactive computer system with two or more displays that are in the same general space (e.g., the same room) and that are related to one another in some way such that they form an overall logical workspace. Since the term *display* has been used to mean many related, but different, concepts, it could cause confusion. I will define *display* for the purpose of this dissertation as an array of light elements that are arranged in a continuous and regular way along the same surface in order to show dynamic output from a computer. My definition is more specific than the most generic definition of display, which includes almost any physical support for graphical information – not necessarily electronic (see, for example [Crabtree et al. 2004]). Although my definition includes displays with curvilinear and irregular forms, most displays are planar and rectangular in shape, and these are the type that I will assume for this work.

Displays can be combined in many different ways to build multi-display environments for different tasks and purposes. In this work I will consider five canonical situations:

- *Multiple monitors at one computer*. This setup is now supported by most commercial graphics cards; two displays is most common, but more are also possible (Figure 4.A).
- *Large composite displays*. These involve multiple identical displays or projected images organized to create a large high-resolution display surface (Figure 4.B). The displays are normally vertical and can use various technologies such as rear projection or plasma screens.
- *Advanced office systems*. These systems combine tabletop displays, wall displays, and standard PC monitors to create a large work surface that is used primarily by a single

person. Different sizes and resolutions can correspond to different task demands (e.g., focused work on a high-resolution monitor vs. peripheral attention on a large projected display – see Figure 4.C).

- *Meeting rooms*. These systems extend the office-of-the-future idea to a primarily collaborative space, where multiple people work at multiple displays. In these environments, large displays are used to enable public presentation and shared focus, and smaller displays allow individual or small-group work (Figure 4.D).
- *Composed workspaces from linked mobiles*. People can form an ad-hoc MDE with their individual mobile computers (laptops or PDAs – Figure 4.E). These environments can either take the form of a contiguous shared workspace, or of a simpler workspace that allows for more limited interactions such as transferring files to others' computers. (Note that both of the preceding environments can also involve mobile devices).

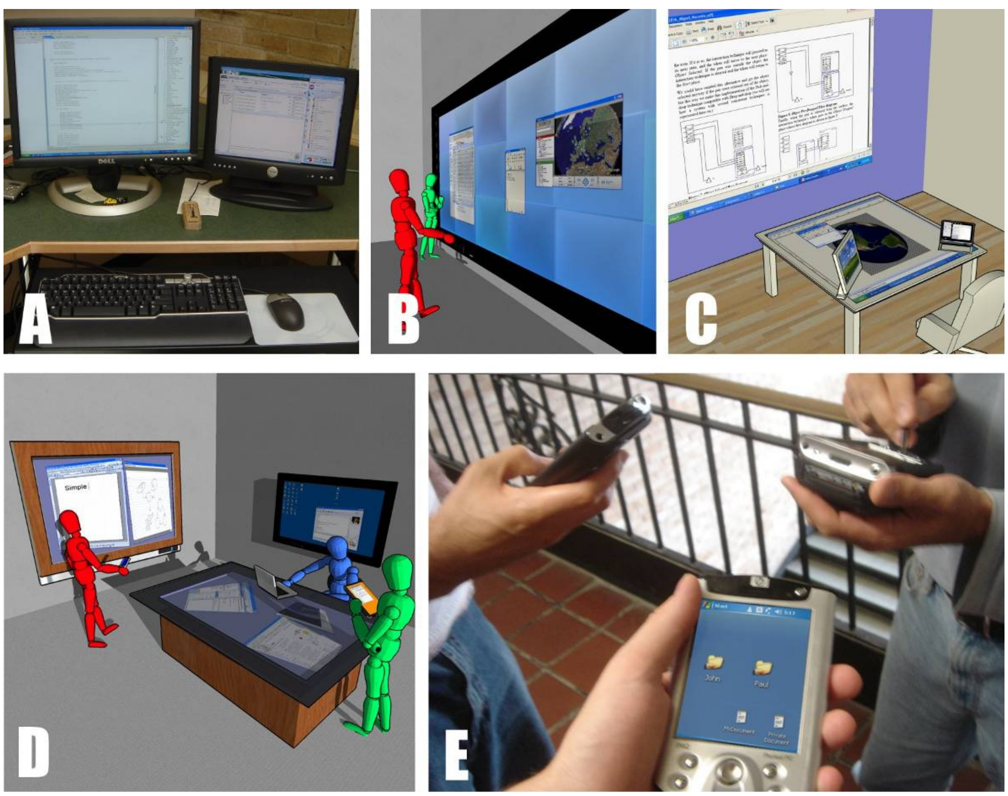

Figure 4. Different kinds of MDEs. A) multi-monitor computer, B) large composite display, C) advanced office system, D) meeting room, E) linked mobile composite environment.

1.3. Problem Statement and Research Hypothesis

The problem addressed by this research is the lack of knowledge about how the characteristics of cross-display interaction techniques affect performance, user preference and subjective workload in cross-display object movement. This lack of knowledge has thus far resulted in several multi-display interfaces that are unnecessarily slow, hard to use, and disliked by the users, and has the potential to hinder future development and use of future MDEs.

To address this problem this dissertation will test the hypothesis that *human performance in cross-display object movement actions is affected by the way that the cross-display interaction technique refers to destination displays, by the way that the interaction technique maps the physical environment, and by the way that displayless space is dealt with during the execution of the movement.*

I investigate and test this hypothesis through five laboratory experiments that provide empirical evidence on the performance differences between cross-display object movement techniques.

1.4. Contributions

This dissertation makes three main contributions: the identification of three factors that can significantly affect human performance in cross-display visual object movement action (how displays are referred to, how the physical environment is mapped, and how displayless space is dealt with); the empirical evidence of how these three factors affects performance in CDOM tasks; and an exploration of the design space according to these three factors.

As minor contributions I also provide a taxonomy of existing CDOM techniques, a set of recommendations for the design of CDOM techniques in MDEs, a new mouse-based interaction technique for cross-display cursor interaction (Perspective Cursor), an adaptation of Halo (an off-screen feedback technique) for cursor movement between displays, an adaptation of Halo for non-planar environments, and a working theoretical model for the analysis of cross-display object movement processes.

1.5. Overview of the Dissertation

Chapter 2 introduces and constrains the topic of the dissertation: cross-display object movement, and cross-display object movement interaction techniques; the chapter includes a general literature review of work related to MDEs. Chapters 3, 4, 5 and 6 analyze each of the three sub-problems of CDOM: referring to destination displays, finding the right control action for the right movement, and the execution of the control action (the analysis of the second topic, the mapping of the physical environment, is divided between chapters 4 and 5). Each of these chapters contains detailed reviews of previous work as is relevant to the topics of each chapter, and detailed descriptions of the experiments. Most of the literature review of this dissertation is distributed among Chapters 3, 4, and 6 because related work in each of the three sub-topics is fairly independent to the other sub-topics, and because organizing literature can be better contextualized by being located near to where it is used.

Chapter 7 discusses the findings presented in the three preceding chapters in a general context. Chapter 8 concludes the dissertation by summarizing the findings, the contributions and suggesting avenues for future research.

CHAPTER 2: CROSS-DISPLAY OBJECT MOVEMENT

Cross-display object movement (CDOM) is the action by which a user translates the visual representation of a digital object (e.g., an interface window, an icon, a picture, a set of pixels, or the cursor) from one display to a different display in a co-located multi-display environment. In order to accomplish cross-display object movement the user has to communicate with the system to perform the action through a *cross-display object movement interaction technique*.

Although all cross-display interaction techniques enable the same basic goal, they might be different in the details that they allow. For example, some techniques support the repositioning of an object to a particular zone of a display, whereas some techniques just move the object to another display – but not necessarily to a specific location in that display; some techniques allow for moving objects to any display in the system, whereas others are limited to displays physically accessible by the user (range). In the rest of this dissertation I will refer to these capabilities whenever they are relevant with the term *power of a CDOM technique*.

2.1. Scope

To make the topic manageable I have focused this dissertation on the human-computer interaction side of system design. There are many other important issues in the design of cross-display interaction that fall out of this category (e.g., the protocol details of the transfer mechanism between displays, the distributed architecture of multi-display systems, the location where the data is actually stored, the scalability of different implementations), and some of these issues could impact the quality of the user interface. However, my research assumes that implementation issues have been solved or are likely to be solved in the near future. Nevertheless, Section 2.2 contributes a brief review of work on the implementation of multi-display environments and on the input technologies that interaction techniques might rely on.

Similarly, this dissertation does not contemplate the motivation, social processes, applications, or larger tasks that might trigger cross-display object movement; instead the analysis starts at the point when a user of a MDE already knows where an object has to be translated, and looks at how the goal can be achieved through system support. The dissertation is therefore constrained to the level of analysis described in Figure 5: it does not look at either the specifics of system implementation or architecture, or at the motivations that lead to cross-display object movement.

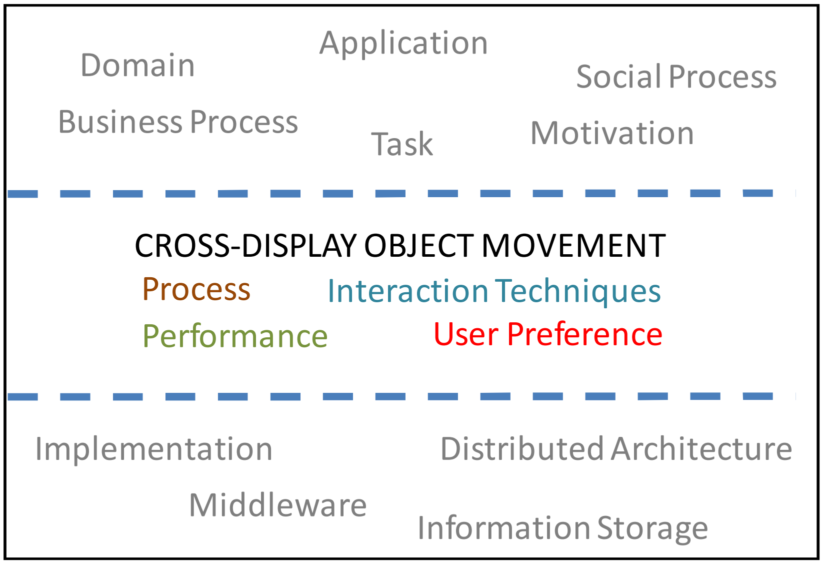

Figure 5. Level of analysis for the dissertation (top is more abstract, bottom is more specific)

The scope of this work is also restricted in that it measures human performance of the cross-display object movement action (i.e., at how fast and with how many errors the task can be accomplished), subjective user preference, and subjective workload. Although there are many other variables that can be measured (e.g., subjective perceived workload, fun, computational complexity, cost), I chose human performance as the most useful measure because it is relatively stable, because it is often correlated to other important measures (if a technique is faster, it seems likely that it requires less effort and it will be preferred by users[2]), and because it is the dominant measure for the evaluation of interaction techniques in current and past literature (e.g., [Balakrishnan, 2004, Benko & Feiner, 2005, Bezerianos & Balakrishnan, 2005, Cockburn & Firth, 2003, Grossman & Balakrishnan, 2005]).

Finally, all interaction techniques were evaluated using adult men and women that had previous experience with computers and standard graphical user interfaces (GUIs), and without any relevant sensory, motor, or mental disabilities. Although I believe many of the results and

[2] The experiments reported in Chapters 3, and 5 provide some evidence of the correlation between performance and user preference: the fastest techniques are also the best liked techniques.

conclusions of this work are applicable to populations beyond the ones tested (e.g., children, older adults, persons with disabilities), this will require further testing that falls out of the scope of this dissertation.

2.2. Technologies that Enable CDOM

Cross-display interaction techniques depend on a number of underlying technologies; in this subsection I step back to set this research in the wider context of system design. The following sections are brief surveys of the hardware and software infrastructures and the input devices required to implement cross-display interaction techniques. Note that many technologies and architectures can support cross-display object movement techniques; however, a full review of networking architectures and input sensing would require much time and space, and is out of the scope of this dissertation. Instead, I provide a limited review of technologies that are currently used or have been investigated in the context of multi-display interaction.

2.2.1. Hardware and Software Infrastructure

The simplest kind of MDE in use today is a standard computer that is equipped with a dual-output video card and two monitors. This setup is supported by most current operating systems [Apple 2008, Microsoft 2008, Wikipedia 2008], and is no more difficult to program than single-display machines; the display space, although divided into two different display surfaces, is still considered a single virtual space by the operating system and allows seamless transitions of objects from one display to the other.

This approach can be extended by adding extra video cards to a single machine. Building complex MDEs in this manner has limits, however, because not all kinds of displays can be connected directly to the machine through a cable (e.g., mobile displays), and because a single machine might not be able to deal with all the input, computation, and output required by the users of a complex MDE.

Some researchers have instead proposed the use of meta-operating systems (also called middleware infrastructures) that combine a number of different machines connected through a network into an integrated interface (e.g., i-Ros [Ponnekanti et al. 2003], Gaia [Román et al. 2002], and Beach [Tandler 2002]). Meta-operating systems also take care of important issues in the exchange of objects such as where a transferred object really resides (i.e., is the object instantly copied into the destination device's memory, is it kept in a central repository, or it is merely an iconic representation that is transferred?), and what are the access permissions for objects in other displays. These issues are important in multi-user systems because they can affect privacy and safety of the data exchanged between displays and devices.

Meta-operating systems are powerful and can greatly simplify the implementation of cross-display interaction techniques. However, there exist easier ways to connect two or more devices (and their displays): for example, if two devices are connected through a network, it is relatively easy to support data transfer using e-mail, instant messaging, or other basic transmission services. However, for these kinds of systems the problem is the configuration and specification of the connection through which two devices are to exchange data, because it sometimes requires the selection of the devices among a list of all possible other devices in the network. This problem is critical for the success of MDEs and has been addressed by research such as SyncTap [Rekimoto et al., 2003], Smart-Its friends [Holmquist et al., 2001] and Bump [Hinckley, 2003], which reduces configuration and specification processes to natural gestures.

Although the underlying hardware and software implementations of an MDE can affect the viability and performance of cross-display movement techniques, I will henceforth assume that the necessary underlying software and hardware are present and that they meet the requirements of the interaction technique at hand (e.g., interaction speed, efficiency of data transfer). For most interaction techniques I assume as well that a working connection between the involved devices has already been established.

2.2.2. Input Device Technology

This dissertation deals with many interaction techniques that are of several different types. These techniques rely on a broad set of input technologies which can affect the performance, usability and monetary cost of the designed interfaces. Most of these technologies are widely used and do not present problems for the implementation of the techniques presented in this work: for example, indirect devices (e.g., mouse or trackball [MacKenzie et al., 2001]), touch screens [Sears & Schneiderman, 1991], pen-based devices [Wacom, 2008], or buttons and keyboards [Li et al., 2005] (see [Hinckley, 2006] for a current survey).

Other interaction techniques require input that goes beyond traditional input devices: for example, some techniques often require tracking of people and devices in the shared space (e.g., [Nacenta et al. 2007b]). The problem of tracking objects and people has not yet been completely solved; however, multiple research groups are very active in this area. Recent projects make use of a large array of different technologies to achieve different levels of accuracy, ranges, and costs in 2D and 3D tracking; for example, Hazas et al. [2005] and Harter et al. [1999] use ultrasound, Randall et al. [2007] use light measurement, Vildjiounaite et al. [2002] use a combination of accelerometers, magnetic sensors and map knowledge, Spratt [2003] uses signal information from wireless links, and Ji et al. [2006] use WiFi signals (see the overview by Manesis and Avouris [2005]).

A discussion of the advantages and disadvantages of each of these technologies and systems is out of the scope of this dissertation. Instead, I assume that the technology underlying the techniques discussed below is already reasonably well implemented or will be in the near future.

2.3. Research on Multi-display Environments

This section examines previous literature on MDEs that is not necessarily related to cross-display object movement. Reviews of previous research on cross-display object movement and CDOM interaction techniques appears in Sections 3.1 and 6.1 and Chapter 4 as it relates to the respective topics of the main topics of this dissertation.

2.3.1. Human Performance with MDEs

Multiple research groups have studied how single users operate MDE interfaces. For example, Grudin [2001] observed how monitors in multi-monitor desktops are either primary or secondary, and how these are used for different purposes; Hutchings and Stasko [2004], and Hutchings and colleagues [2004] looked at display space management in MDEs; Tan and Czerwinski [2003] studied the effect of different depths and sizes of displays in visual tasks; Ringel Morris [2003] looked at how MDEs compare to multiple virtual desktop on a single display; and MacIntyre and colleagues studied how to use large secondary displays for peripheral information [2001].

The perceptual peculiarities of different displays have also been studied. For example, Wigdor and colleagues [2007] measured the effects of tilt on the perception of visual variables; we studied the perception and interaction with perspective-distorted and perspective-corrected data [Nacenta et al., 2007b; Hancock et al., 2009]; and Hancock and Carpendale [2007] have looked at how best to interact with 3D objects in flat horizontal surfaces.

2.3.2. Interaction Techniques for MDEs

There are also interaction techniques specifically designed for MDEs that are not directly related to CDOM. For example, Mackinlay and Heer [2004] designed techniques to mitigate the visual effects of the discontinuity of displays; Tan and colleagues [2004] and Biehl and colleagues [2008] have explored the possibilities of replicating regions of a display in another display; and Mandryk and colleagues [2005] have used a pseudo-haptic approach to take advantage of the edges between monitors in multi-monitor systems. There are also a number of commercial applications that help to deal with mostly technical (but also interface-related) problems of multi-monitor machines [Realtime Soft, 2009; Synergy, 2009]. However, most of the interaction techniques specifically designed for MDEs are CDOM techniques, and will be reviewed in the pertinent sections of the following chapters.

2.3.3. Collaborative Behavior on MDEs

Multiple studies have looked at how people collaborate in co-located MDEs, and how the different configuration of displays affects collaboration. For example, Inkpen and colleagues [2005] studied a large number of different display factors (angle, size, number, and user arrangement) on a route planning task; Rogers and Lindley [2004] investigated how the orientation of different displays affect collaboration processes (tables vs. wall displays); Rogers and colleagues [2009] compare different input settings for co-located groups in terms of verbal participation and interaction with virtual objects; and Haué and Dillenbourg [2009] look at how the number of individual screens (laptops) can change collaborative processes.

2.3.4. Smart Rooms and Other Multi-display Settings

Most other research on MDEs has been focused on the design of novel smart rooms that integrate a number of technologies. Initial efforts in the 1980s and 1990s were focused on developing software and hardware that could enhance meetings and improve group processes [Nunamaker et al., 2009; Mantei, 1988; McGrath and Hollingshead, 1993]. This approach has been replaced by the design of systems that are less intrusive and instead try to support the natural collaborative processes of groups through advanced technologies [Johanson et al. 2002, Tandler, 2002, Tandler et al., 2001, Holmer et al., 1998]. However, the interfaces of the later systems are often more a product of the technology and its limitations than of specific interaction design.

Besides the full general-purpose systems discussed above, there has been research on particular applications that take advantage of several displays. For example, Wigdor and colleagues [2006] investigated control room applications; Robertson and colleagues [1996] built a prototype that combined PDAs and a television set for the browsing of real estate; Forlines and colleagues [2006] implemented a geographic visualization system based on Google Earth; Wigdor and colleagues [2009] designed, implemented and evaluated a system for scientific collaboration; Lanir, Booth and Tang [2008] developed a system to allow presenters to take advantage of extra screens, and Rogers & Rodden [2003] have built a number of multi-display setups for different purposes, such as travel planning for clients and travel agents.

2.4. The Process of Cross-display Object Movement

Investigating the effects on performance of cross-display interaction techniques (the main goal of this dissertation) requires a characterization of the design space of cross-display interaction techniques. The characteristics in which interaction techniques differ from each other can then be used to explain differences in performance. In other words, in order to understand and observe differences in human performance between techniques of type A and type B, it is necessary to establish the elements of interest that classify interaction techniques into either group.

Many different characterizations of interaction techniques are possible; for example, we can characterize interaction techniques according to their input type (direct, indirect, absolute or relative input), mode (visual, auditory, haptic), visual representation, monetary cost or many others.

I chose to investigate technique characteristics derived from a simple cognitive procedural model of the action of moving visual objects across displays. I look ‘under the hood’ of the general process of CDOM to find different sub-processes, each of which can be affected in different ways by specific characteristics (design decisions) of the interaction technique.

The advantages of this approach are that it allows for a choice of defining characteristics that are intrinsic to the action of interest (and therefore unaffected by external factors such as implementation details), and that it allows us to focus on characteristics that are specific to multi-display environments (i.e., it is mostly orthogonal to other characteristics that might affect performance on any platform – not necessarily multi-display – such as the directness of input, the particular input device, etc.) The characteristics that result from this approach also have the advantage that they translate directly into design decisions in the design space of CDOM techniques.

The simple model is composed of two basic stages on the human side: *action planning* and *action execution* (see Figure 6). It starts modeling CDOM at the point when a user has already decided that she wants to move an object and knows to which display (or particular destination) the object is to be moved. Action planning is informed by the environment (task, intentions), and, in turn, informs action execution, which usually happens through a continuous interaction with the system interface. Note that, although action planning is a mental process of the user that is not directly affected by the system itself, the design of the interaction technique (once its basic use is learnt) will condition how users plan their actions. In other words, users plan CDOM actions according to their knowledge of how the interface of the system works[3].

The following subsections describe the sub-processes of the model, explain how the relevant characteristics of interaction techniques that affect each process are derived, and map the experiments to the conceptual structure of this dissertation (see Figure 6 and Figure 7).

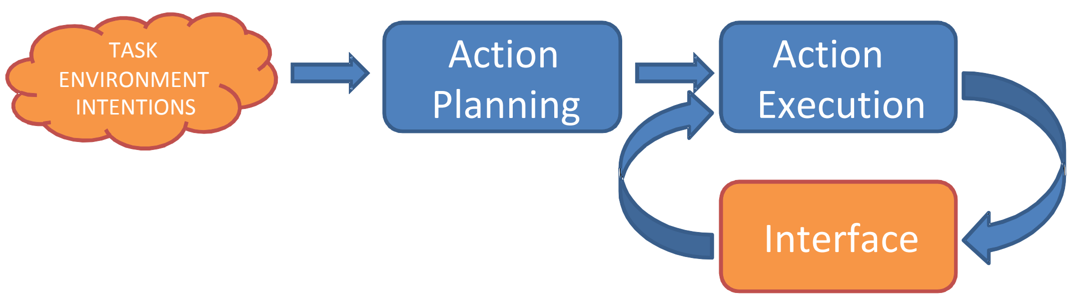

Figure 6. Simple model of the stages of CDOM. Blue elements (action planning and Gesture Execution) involve the human user whereas orange elements represent non-human elements (the external environment and the system interface).

[3] In this dissertation I will not deal with learning or discoverability aspects of interaction techniques. The focus is on the performance of techniques once they have been learnt, which is the overwhelming majority of the time of use because CDOM techniques are typically learnt in a few trials.

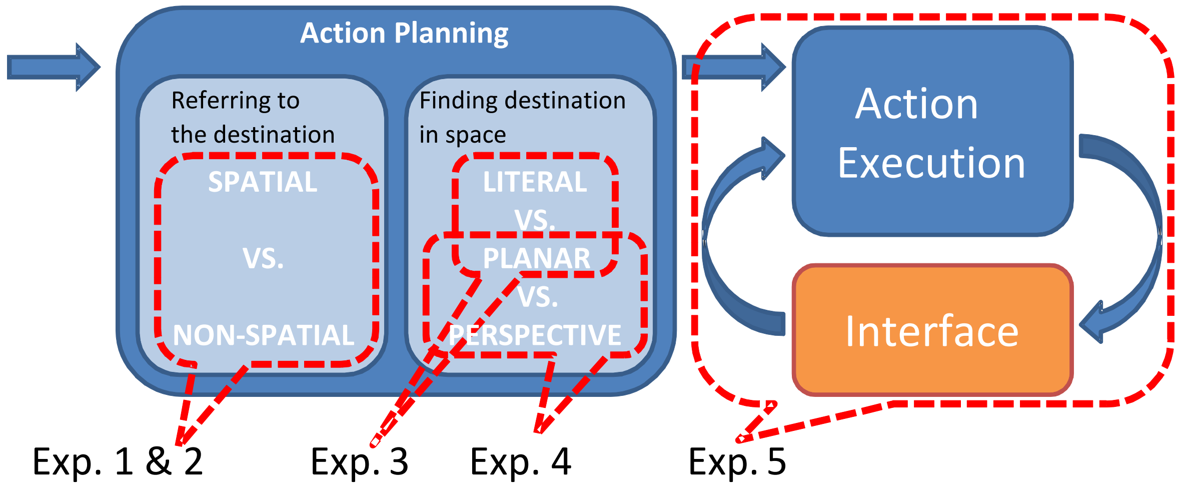

Figure 7. An extended version of the model in Figure 5 that maps the experiments into the conceptual structure of the dissertation.

2.4.1. Action Planning

The first sub-process is the planning of the action, that is, the internal process by which the user translates the intention (e.g., I want to move object X from display A to display B) into a concrete plan of action (e.g., I have to pick it up with my pen from display A and then touch display B). The interaction techniques that are available (and known by the user) determine how the action planning is carried out. For example, it is not the same to plan a Pick-and-Drop action that will involve touching the object and then the destination display with a pen [Rekimoto, 1997], as it is to choose the destination display from a list (as in Multibrowsing [Johanson et al., 2001]). With Pick-and-Drop the plan will have to include how to physically reach the object and the destination; in Multibrowsing, the plan involves identifying the destination display's name in the list and clicking on it.

Within action planning I consider two elements that can affect performance: how the destination is referred to (i.e., how the user indicates to the system which display should receive the object) and, in case that the displays are referred to spatially, how the interaction technique control movement relates to the actual physical space around the user.

2.4.1.1. Referring to the Destination

At some point during the process of cross-display object movement the user must somehow indicate to the system which display (or location within a display) is the destination. There are many ways to refer to a display; the example in the paragraph above already shows two different ways to refer to displays (by name, or by touching). Chapter 3 analyzes how techniques that require spatial or non-spatial ways of referring to the destination will affect performance; the chapter contains a discussion of previous empirical and theoretical work, and then describes two experiments that compare spatial and non-spatial techniques in two different situations. Experiment 1 compares the performance of techniques for a spatial CDOM task, i.e., a task where the destination of the object is formulated in a spatial way. Experiment 2 compares the same techniques for a task where the position of the destination is not relevant.

2.4.1.2. Finding the Right Control Action for the Destination

If the display is referred to spatially (e.g., if the interaction technique requires users to point with a finger at the destination display, to move the mouse in that direction, to select the display from a miniature map, or to directly touch the destination) there is another consideration that might affect how fast and accurately the action is planned: what is the relation between the required control action (e.g., moving the mouse), and the actual environment surrounding the user? Chapters 4 and 5 investigate how performance is affected by the different types of mappings that an interaction technique can make between input and the physical environment. Chapter 4 contains the main theoretical approaches and a discussion of the main alternatives. Chapter 5 reports two experiments that compare three different types of techniques that are based in three different mappings: planar techniques bring the different display input spaces together as if they were laid out on a virtual flat surface; perspective techniques use a mapping that adapts dynamically to the position or perspective of the user; and literal techniques map input and physical environment in a direct fashion (input overlaps the physical environment one-to-one). Experiment 3 compares planar techniques with literal, and Experiment 4 looks at planar and perspective techniques.

2.4.2. Movement Execution

Once the user has planned the action (or sometime after having started to plan it), the user has to actually execute the action or command that will perform the actual object movement. There are several important elements that affect how the execution of the action is carried out. The most general issue that affects MDEs is the presence of feedback during execution. The execution of the action can happen with the aid of a feedback channel that continuously indicates the user's progress, and allows her to correct the gesture along the way to achieve the desired result. For example, the laser pointer technique [Oh & Stuerzlinger, 2002] projects a red dot indicating the current indicated destination through the whole action. However, the interaction technique might not implement this back channel (e.g., the flick technique [Reetz et al., 2006], in which the action happens almost instantaneously, without the opportunity to change the result once the flick gesture is executed). The presence or lack of feedback during all or part of the execution of a gesture is particularly relevant to MDEs because these are, by nature, fractured: displays can provide visual feedback, but the space in between cannot. Chapter 6 discusses feedback in detail and analyzes previous research from several fields that relate to this issue. The experimental work of the chapter looks specifically at displayless space, an issue that, although directly affected by the presence of feedback, also involves a number of other factors such as the consistency between input and visual information and the effect on performance of extended motor movements. The results of the empirical investigation of Chapter 6 have direct consequences on design for generic as well as existing MDEs such as multi-monitor systems.

2.5. Research Methodology

The main research method used in this work is the controlled laboratory experiment. The following four chapters present experiments that investigate hypotheses on how design decisions affect each of the three sub-processes in which I divided the CDOM action (Chapter 4 does not describe any experiment, but instead analyzes the relevant previous literature and technique types for the experiments in Chapter 5).

There are many other methods that could be used to investigate the design space of cross-display object movement; however, controlled laboratory experiments are arguably the best method to test the hypotheses, given that I focus on performance and that this is the first systematic exploration of this design space. Further research will necessarily require other methods such as observations in the field, contextual inquiries, or designer explorations, but this falls out of the scope and size of this dissertation.

Each chapter contains a full review of interaction techniques that are related to the particular sub-process, a summary of relevant literature (including literature outside of the HCI field), a full description of each of the experiments, an extended discussion on the results of each experiment and the set of experiments as a whole, and the implications for design. Chapter 7 summarizes the results from the previous three chapters, and discusses them as a whole. Chapter 8 presents the main conclusions.

CHAPTER 3: MOVEMENT PLANNING - REFERRING TO THE DESTINATION

One of the most important pieces of information to be communicated from the user to the system in a cross-display object movement is the destination display. This information has to be encoded by the user in a way that is determined by the CDOM interaction technique. The possible encodings required by the interaction technique can be very different, from the listed names of available destination displays shown in Multibrowsing [Johanson et al. 1999] to the directional pointing of a destination display used in the classic Put-that-There technique [Bolt, 1980]. Similarly, people can encode the intended destination display in different ways; for example, a user might want to move an object to “this display on my table” or to “Janice’s computer” (regardless of where the display is). This chapter studies how the way in which interaction techniques and users refer to destination displays (the “encoding”) affects performance. The focus is now on the leftmost box of the diagram in Figure 7, which represents the cognitive process of referring to a destination display.

3.1. Background

3.1.1. Spatial vs. Non-spatial Techniques

The characteristic that I use to differentiate how CDOM interaction techniques refer to destination displays is *spatiality*. An interaction technique is spatial when the action required by the user encodes spatial information of the location of the destination display in the environment. For example, consider the early Put-that-There technique [Bolt, 1980]. In this technique, the user points with the finger (or with some other device) to indicate the intended destination (see Figure 8). Note that although the Put-that-There technique uses voice to activate the selection and release of the object, the information about where to put it is specified only by the pointing gesture.

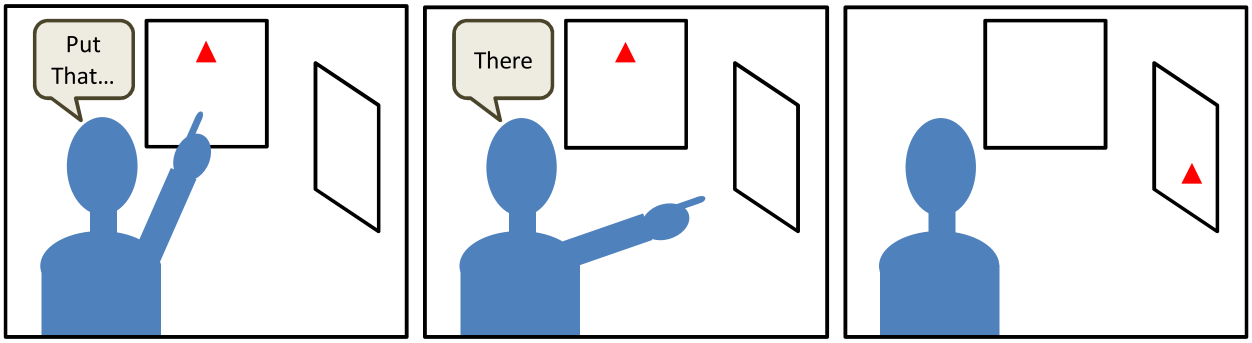

Figure 8. The Put-that-There technique as envisioned by Bolt [1980] (from left to right).

Most existing CDOM techniques are spatial, although there are several different varieties. Laser pointing techniques [Oh & Stuerzlinger, 2002, Voida et al., 2005, Myers et al., 2002, Parker et al. 2005] (also called remote pointing, laser beam, or distance pointing techniques) belong in the same group than Put-that-There. Variants usually differ in the device used to point and the activation method (e.g., voice, a button integrated in the device, a button in the other hand, a hand gesture, and dwell time), but are essentially identical in how the direction of the arm (or the device) indicates the destination.

Direct input techniques such as Pick-and-Drop [Rekimoto, 1997] (See Figure 9), or Passage [Prante et al., 2004] also make use of spatial referencing, but these techniques require direct physical contact with the destination display. Other examples of spatial interaction techniques are those based on the world-in-miniature paradigm (WIM – graphical representations of the physical space) [Bieh and Bailey, 2004, 2006; Swaminathan and Sato, 1997] and those in which the cursor, steered by a mouse or any other indirect input device, jumps from one display to another according to the physical location of the displays [Wallace et al., 2008; Johanson et al., 2002; Baudisch et al., 2004; Benko and Feiner, 2005, 2007]. Notice that even though in this last group the spatial relationship between the gesture and the physical space is indirect, the gesture still relates to the spatial arrangement of the physical space. For example, in world-in-miniature techniques, the gesture relates to the location on the miniature (map), and this in turn, is a representation of the physical reality. If the relationship between the map and the physical layout of the MDE is not spatial (e.g., if the different displays are arranged in the map according to their frequency of use, their owner, etc.), then the technique cannot be considered spatial, but would fit in the non-spatial category described below.

This brief summary of spatial techniques should give an idea of the variety possible within spatial techniques. Although there are more existing and possible spatial techniques, I postpone its discussion to Chapters 4, 5 and 6, which deal in more detail almost exclusively with differences in performance of spatial techniques of different types. This chapter, however, is focused on differences due to how techniques reference to displays.

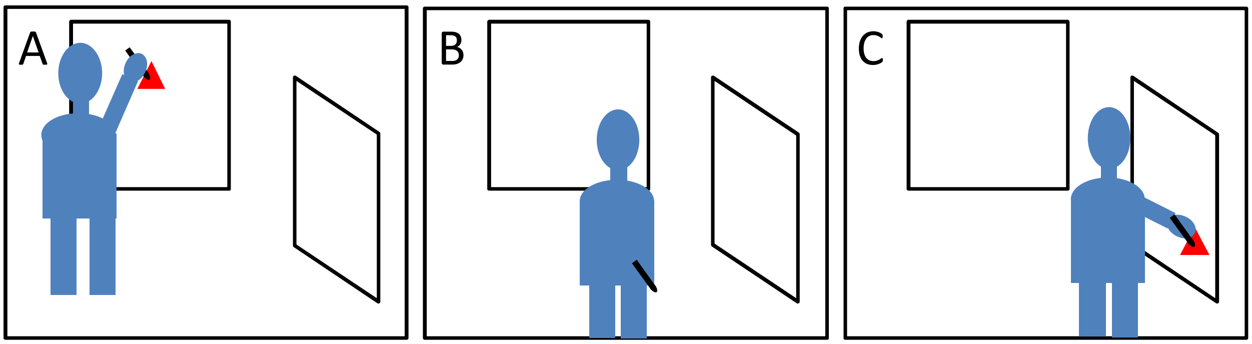

Figure 9. The Pick-and-Drop sequence. A) The user touches the object with a pen and clicks a button. B) The user moves to the destination location (while holding the button). C) The user releases the button while touching on the desired destination.

CDOM interaction techniques where the required action by the user does not relate to the spatial layout of the displays in the physical space are called *non-spatial* techniques. The most common examples of non-spatial techniques are those that use some kind of naming of the displays. For example, in Multibrowsing [Johanson et al., 2001], the user can redirect content from one screen to another by selecting the destination from a list of available displays (see Figure 1 in page 3). In this case the interaction technique is non-spatial because the control movement (moving the mouse so that the cursor is on top of the correct item of the list) is not related to where the display is in physical space. Instead, the destination display is indicated by a symbol (the name of the display).

The same kind of referencing by name is common in cross-device file transfer techniques such as Instant Messaging (IM) [Microsoft, 2008c], Shared Folders [Microsoft, 2008b], or E-mail[4]. For example, in IM, instead of selecting the destination display by the name of the display (as in Multibrowsing), the user has to select the name of the buddy that is connected to that

[4] Instant messaging, shared folders and e-mail were not originally intended as cross-display object movement techniques but they are often used for that purpose in co-located situations due to the lack of more specific support for CDOM. For example, it is common to pass a file through IM to a nearby co-worker because the two machines are usually connected through a LAN or the Internet, but not connected at the level of an interface that knows of their physical proximity.

display from the buddy list. E-mail can require a very similar action (selecting from an address list) or an even more symbol-based action, such as typing the address of the recipient. In contrast, shared folders can be more spatial in the sense that the icons of the remote folders can be spatially distributed around the desktop. In fact, if the user takes the time to arrange them according to the physical arrangement of their destinations, this technique can be considered spatial (it becomes a map of sorts – see Figure 10). However, if the shared folder icons are distributed in the display randomly, the technique is equivalent to IM and E-mail described above, with the exception that the input movement is not limited to one dimension as in selecting from a list, but it is instead two-dimensional (the control movement can go in any direction within the plane).

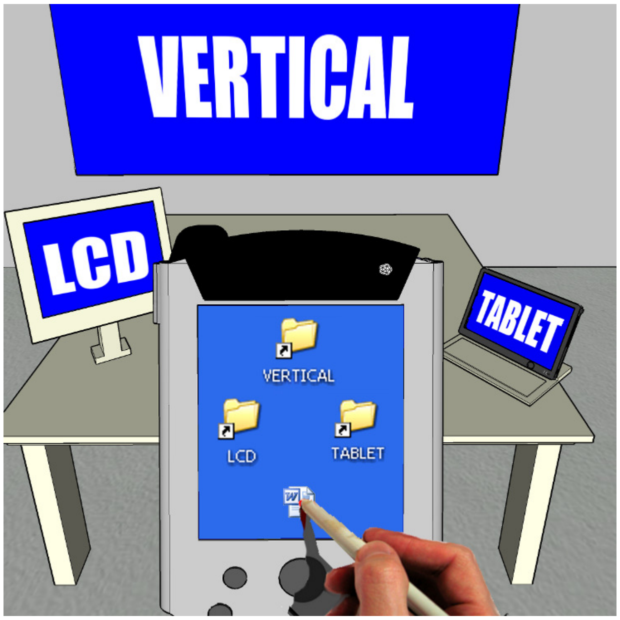

Figure 10. An example of Spatially Arranged Folders (a spatial technique).

There are a large number of potential variations of techniques based on symbolic selection of the destination. For example, symbolic links could be established between a destination display and color, shapes, ideograms, hieroglyphs, photos of the display, sounds, tactile textures, and almost anything imaginable. However, most of these do not make sense and would be difficult to use. Most existing techniques use text (i.e., the name of the display).

Aside from symbolic techniques, there exist other non-spatial techniques that are based on the navigation of some data structure related to the destinations. Two examples of these are the keyboard-switch Multi-Monitor Mouse and the mouse-button-switch Multi-Monitor Mouse presented by Benko and Feiner [2005, 2007] (most other techniques presented in this paper can be classified as spatial). In these variants of the Multi-Monitor Mouse, the cursor switches from one screen to the next in a circular list of available displays, much as the Alt-Tab key combination switches between applications in Windows™ systems.

Unfortunately, it is difficult to further categorize techniques based on non-spatial referential domains or to study them all in detail because of the countless possibilities for assigning symbols or descriptions to displays or locations inside displays (symbolic techniques) and because of the large number of possible ways to traverse the space in navigational techniques. For the rest of the chapter I have chosen text-based symbolic techniques (e.g., naming schemes such as the one in Figure 10, or IM) as the most prominent representative of the non-spatial group of techniques. Although other non-spatial techniques can offer an unexplored and potentially fruitful direction of study (one that certainly deserves further research), I do not have any reason to believe that other symbolic ways to refer to displays (e.g., color, texture, etc.) can outperform text.

3.1.2. User Encoding and Interaction Technique

Just as interaction techniques require some kind of encoding that refers to a particular destination, people also encode destinations differently when formulating their intention. Consider, for example, the difference between trying to move a digital object to "the large display in front of me" as opposed to "my PDA". The difference is that the latter formulation does not depend on the location of my PDA (it can be stored in my pocket or behind another display), whereas the former formulation depends completely on the physical spatial layout. In other words, users also encode the destination display in spatial and non-spatial ways. Figure 11 shows examples of encodings used by techniques and people when performing cross-display object movement tasks.

The type of encoding of the destination information in the user's mind depends mostly on the type of task that originated the object movement. Using the same examples above, if a user wants to move a document to her personal digital assistant (PDA) to inspect it while she is on the go, she probably does not care about the exact current location of the PDA (it might even be already in her pocket), but she does care that it goes to a PDA (that she can take with her) and that it is her PDA (the document will be no use to her if it is in somebody else's PDA). However, if what the user wants is to move the object to a large display in front of her in order to inspect the data of the document, the relevant information about the destination is where it is located (in front of her). Therefore the task will likely determine if the user encoding is spatial or not.

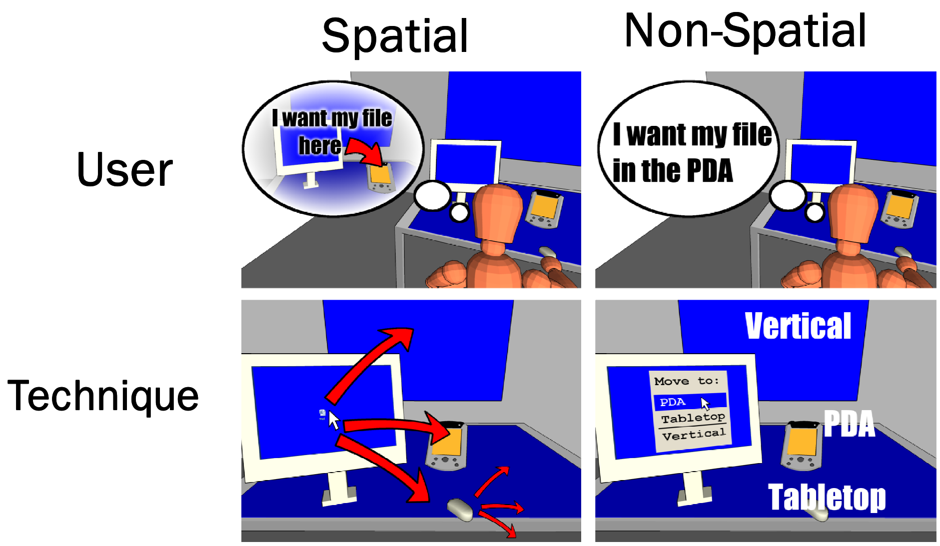

Figure 11. Examples of the encodings of the user and the technique (top and bottom rows respectively) according to spatiality (spatial – left, non-spatial – right).

The main working hypothesis for this chapter is that performance (how quickly and accurately users can move an object to another display) depends on the match between the user encoding of the destination and the encoding required by the interaction technique (i.e., the way the user and the interaction technique refer to destinations). In particular, it is expected that spatial interaction techniques will work better than non-spatial techniques for tasks where the destination is thought of in spatial terms. The underlying assumption is that if there is no match between interaction technique and task, the encoding of the task requires an extra mapping process that will take extra time and can be prone to error, resulting in worse overall performance.

Another interesting question is whether the co-located nature of the scenario can also affect performance even if the task and the interaction technique are non-spatial. In other words, is it possible that, because of the spatial context in which the task takes place (a co-located setting), non-spatial techniques are in disadvantage even if the task is formulated in non-spatial terms?

The following section reviews existing research that is relevant to these issues.

3.1.3. Previous Studies

Biehl and Bailey [2006] compared two spatial application relocation methods to what they call a 'textual interface'. The task was a cooperative task that required the movement of applications between large shared displays and personal tablets in order to compose a collage or a comic strip. They found that relocation times were several times shorter with the spatial interfaces than with the interface that used name references for displays. From the description of the task we can deduce that the user's window movements were mostly motivated spatially (e.g., moving the applications from 'there' to 'here'). This study did not control for the type of encoding that the task required and therefore does not provide conclusive evidence about the effect of a good match between the spatiality of the interaction technique and the task (this was not the main goal of the experiment either). However, it indicates that, for the general task they tested (visual layout), spatial techniques are likely to be superior in performance.

In a laboratory study, Swindells et al. [2002] (see also [Swindels, 2002]) evaluated gesturePen, a pointing-based spatial technique that requires pointing at a certain distance to select a device. They compared the spatial technique to a technique based on the selection from a list of devices (a non-spatial technique). The experiment required that the participant select a device from the physical environment that was verbally indicated by the experimenter (by saying the name of the device, e.g., "the laptop"). In the spatial condition, the participants would have to point to the referred destination with the device, whereas in the alternative condition the participants would have to read a label pasted on the device (the label contained an easy-to-read IP-address of the device – based on vegetable names), and find the label in the list interface. They found that selecting a destination took significantly longer when using the list than when pointing. These results indicate that an extra mapping process (the one necessary to find the label and then match it in the list) does slow down users; however, the experiment was not designed to fairly compare the spatial and non spatial alternatives, since the extra mapping required with the labels is not intrinsic to non-spatial interaction techniques (the labels used could instead have been directly the names of the destination devices).

In three laboratory studies, Jones and Dumais [1986] compared the accuracy of recall of text objects (text articles from a journal) when using filing systems based on spatial, non-spatial and mixed classification. The experimenters asked users to identify the filing record of a particular article (from a short excerpt of it). This study could be considered a test of spatial and non-spatial interaction techniques for a task that requires a non-spatial encoding of the "destination" (in this case, the article). Their results indicate that spatial filing does not have any advantage in recall accuracy. In other words, storing objects associated with a location in space does not seem to result in fewer errors when trying to remember a particular object. Although memory is not the main focus of the current study, it plays a role in the effectiveness of interaction techniques as we will see from the results of Experiment 2 in this chapter.

3.1.4. Stimulus-Response Compatibility and the Dimensional Overlap Model

The *Dimensional Overlap* model [Kornblum et al., 1990; Kornblum, 1992] is a cognitive theoretical model that describes and predicts human performance in stimulus-response (S-R) tasks. Although classic S-R tasks and S-R Compatibility are not directly applicable to the subject of this chapter, I summarize existing work in this area because S-R Compatibility motivated the development of the Dimensional Overlap model, and because the concept of S-R Compatibility will be useful in subsequent chapters.

3.1.4.1. Stimulus-Response Compatibility

A classic *stimulus-response* (S-R) task consists of pushing one of a set of buttons when one of a set of lights is turned on (see Figure 12). Users are asked to react to the onset of the lights as fast as possible according to different mappings. For example, in one condition the button to press is the one immediately below the light (Figure 12.A – also called a *compatible* mapping), whereas in other conditions the mapping has a non-direct spatial correspondence (Figure 12.B and C – two of the many possible *incompatible* mappings).

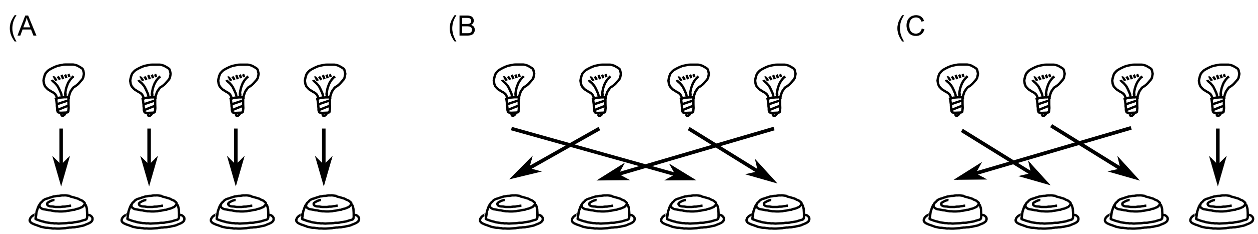

Figure 12. A typical stimulus-response task. The participant has to push a button when one of the lights turns on. Participants are tested with different mappings (A, B, C).

Stimulus-response studies consistently show that when the mapping between stimuli (the lights flashing) and the response (pushing a button) is compatible, reaction times are faster and less error prone, *even after extended training of the subject with the mapping* (see, for example Fitts and Deininger [1954], Fitts and Seeger [1953], Chua et al., [2003], Proctor and Reeve, [1990], Proctor and Vu [2003]).

3.1.4.2. The Dimensional Overlap (DO) Model

Although S-R Compatibility can be used to predict performance in simple graphical tasks, its narrow definition does not help predict the effect of different mappings or encodings of stimuli and responses. For example, what is a compatible mapping between a list of names and the layout of stimuli? The Dimensional Overlap model generalizes S-R Compatibility to include mappings of a more general class. The main assumption of the DO model is that given a stimulus and response set, the fastest reaction time obtainable with optimal mapping is faster if the sets have dimensional overlap than if they do not [Kornblum et al., 1990]. Dimensional overlap is defined in broad terms as the similarity between stimulus and response sets. If the sets are compatible, higher dimensional overlap results in improved performance and accuracy because the translation between stimuli and response occurs through an automatic cognitive process[5]. If the S-R sets are not similar (i.e., dimensional overlap is small), people must carry out a time-consuming search or a rule application, in addition to an inhibition of the results of the automatic process. For example, if the stimulus is textual but the response has to be spatial, there needs to be a search process in which objects in space are identified with respect to their names.

Besides the advantage in performance of overlapping S-R sets, the DO model predicts that the slope of the *reaction time* vs. *number of alternatives* function is reduced when the S-R sets overlap (Kornblum et al., 1990). This means that highly overlapping S-R sets are more scalable than non-overlapping S-R sets.

Although the Dimensional Overlap model is supported by a considerable number of experimental studies, the model has rarely been applied to human-computer interaction problems (with the exceptions of Po et al. [2005] and Proctor and Vu [2003]).

3.2. Research Questions

The design of the experiments in this chapter is informed by results of related previous experiments and by research on the Dimensional Overlap model. However, the experiments are designed with an eye to the design of interaction techniques for real scenarios where several people share the same physical space. Therefore, the main questions that motivate the experiments are:

[5] Automatic cognitive processes are simple processes that happen unconsciously and very fast [Schiffin and Dumais, 1981].

- For tasks that are encoded spatially, does a spatial interaction technique offer performance advantages over a non-spatial interaction technique?
- When the task does not force a spatial encoding, are there any advantages of spatial interaction techniques?

3.3. Experiment 1: Spatial Task

The first experiment is designed to test whether spatial interaction techniques offer any advantage over non-spatial interaction techniques when the task is encoded spatially. The experiment reproduces a scenario in which several people need to exchange documents from their own personal displays. Experiments 1 and 2 were carried out with the same participants, in the same session and in inverse order (Experiment 2 first); I present them in this order because it better follows the logical order of the argument and because it facilitates the explanation of the results.

3.3.1. Apparatus

The experiment was carried out in a large indoor room where four participants were seated facing each other (see Figure 13). Each participant was assigned one portable pen-based computer at the beginning of the session. Several different computers were used for the study: one Hewlett-Packard and one Toshiba tablet PC, and two Sony Vaio ultra-portable computers. The screen resolutions were 800x600 pixels for the two Sony machines, 1024x768 for the HP, and 1280x1024 for the Toshiba. Each computer ran software developed for the study in C#. To compensate differences in size and resolution of the four devices, the application was normalized to appear equal in size to all participants.

The main interface screen (Figure 14) displayed three circles, one for each of the other participants. Each circle contained the real name of one of the other participants.

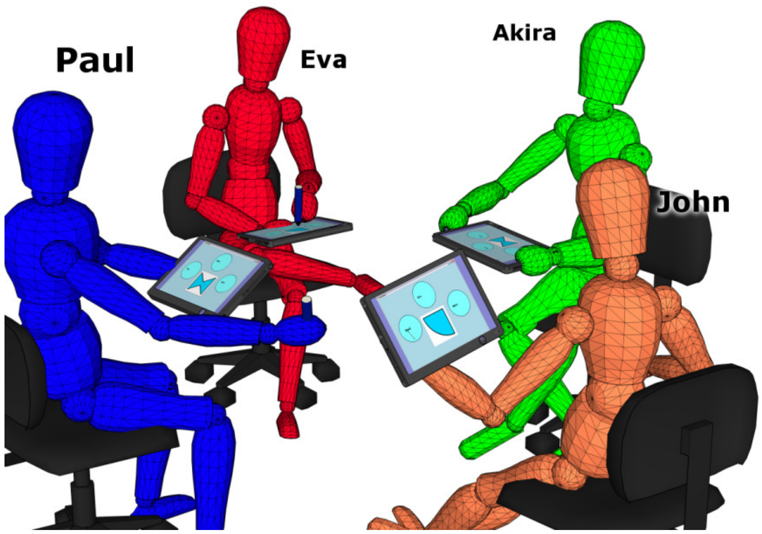

Figure 13. The experimental setting of Experiments 1 and 2.

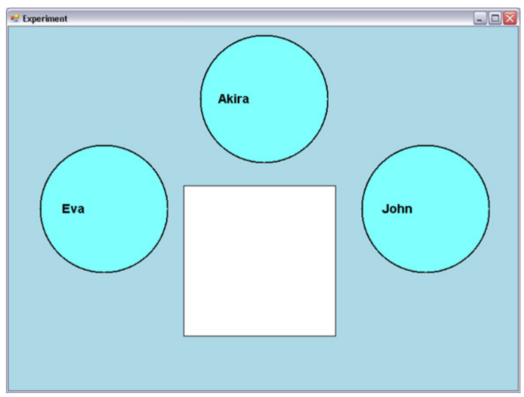

Figure 14. The interface window for Experiment 1.

3.3.2. Participants

We recruited 16 participants from among the staff and student body of the University of Saskatchewan for a compensation of CAN$10. Half were females and half males. The experiments were run in groups of four. Groups were same-sex and all members in the group knew each other from before the experiment. Ages ranged from 18 to 39. The participants read and signed a consent form and a demographic data form at the beginning of the session, and filled in post-study questionnaires after all experimental tasks were completed. Samples of the consent form and questionnaire are reproduced in Appendix A.

3.3.3. Task

Every turn, one of the four participants was selected by the computer and shown a special dialog box that read “me!”. Participants were instructed to say “me!” aloud at the same time that they tapped the dialog. The other participants had then to tap on the center square of Figure 14, and then on the circle that corresponds to the participant that said “me!” as fast as possible and without error. This task reproduces the scenario when one user needs to pass an object to a user who is in a known location. The task represents tasks that require a spatial encoding of the information.

The experimental software in all the devices was synchronized so that only one participant said “me!” per turn, and all participants finished before the next turn started.

The main experimental conditions were *spatial* and *non-spatial*. In the spatial condition, the circles contained the actual names of the participants in the same ordering that the group was seating in the physical space (e.g., Paul – the blue participant in Figure 13 – would see the interface of Figure 14). In the non-spatial conditions the interface was identical, except that the names of the participants did not correspond with their physical locations.

3.3.4. Hypotheses

The experiment was designed to test four hypotheses:

H1: Participants take less time to tap the corresponding circle in the spatial condition

H2: Participants make fewer mistakes in the spatial condition

H3: Participants report less workload for the spatial condition

H4: Participants prefer to perform the task in the spatial condition

3.3.5. Experimental Design

The experiment followed a within-subject repeated-measures design with spatial layout (spatial or non-spatial) as the main factor. For each condition participants performed a short demo block of the task (3 correct trials per person) and then two blocks of 45 correct trials per person; this meant that the group performed 60 correct trials per block, since a trial in which a user had to say “me!” was not counted or measured for that user. The order in which the participants were selected by the system to be the destination did not follow any recognizable pattern. After each block of trials the participants switched seats.

Half of the groups were randomly chosen to do the spatial tasks first; the other half did them in the opposite order. If any participant made a mistake, that trial was replaced for another at the end of the block, assuring that all participants performed at least 45 correct trials each in each block, and a variable number of wrong trials. Once all blocks were completed, they were asked to complete a NASA TLX workload self-assessment questionnaire [Hart & Staveland, 1988], and to select their preferred technique.

We took four measures: completion time (from stimulus onset), errors, subjective workload assessments, and preference.

3.3.6. Results

The completion time measures were first log-transformed to comply with the normality assumption of the data analysis. A one-way repeated-measures ANOVA of the time from stimulus (i.e., the "Me!" button was pressed by the corresponding participant) showed that it took participants less time to complete trials in the spatial condition ($F_{1,15} = 69.29$, $p < .001$), therefore supporting H1. The average time was 711ms in the non-spatial condition and 551ms in the spatial condition, a difference of 160ms (i.e., the task took 29% more time to complete in the non-spatial condition – see Figure 15 and Table 1).

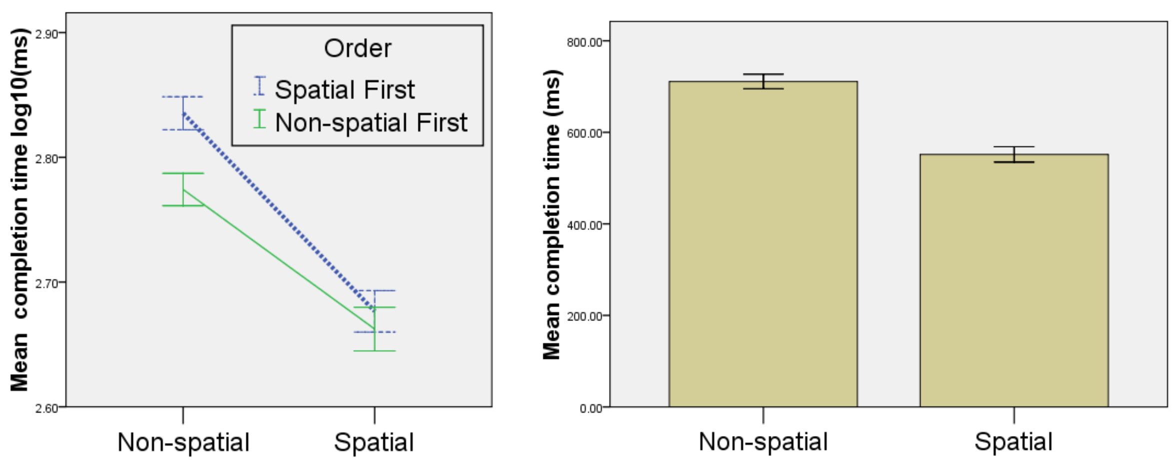

Figure 15. Left) Log-transformed completion time by condition and order. Right) Completion time. Error bars indicate 95% confidence intervals.

Table 1. Average completion time from the "me" signal (measured in milliseconds) and standard error (between parentheses) for the different conditions (columns) according to the order of presentation of the conditions (rows).

	Spatial (SE)	Non-Spatial (SE)	Log10 Spatial (SE)	Log10 Non-Spatial
Spatial First	557.67 (10.86)	769.80 (13.91)	2.67 (.008)	2.83(.007)
Non-spatial First	547.46 (13.17)	658.32 (8.82)	2.66 (.008)	2.77(.006)

The data was tested for order effects with a repeated-measures ANOVA test with condition as within-subjects factor, and order as a between-subjects factor. Order was non-significant ($F_{1,14}$ =

0.011, $p > .42$) and neither was the interaction of order and spatial condition ($F_{1,14} = 2.28$, $p > .15$).

A Wilcoxon Signed Ranks test shows that the number of errors is higher in the non-spatial condition than in the spatial condition ($z = -3.08$, $p < .002$). On average, a participant would make 2.75 errors per block in the non-spatial condition and 0.81 in the spatial. A block is defined as the group of trials that take place without a change of position of the participants relative to each other and the mapping in their devices. This result supports H2.

A Wilcoxon Signed Ranks test was applied to the workload data, revealing a significant difference between conditions ($z = -3.50$, $p < .001$), with the non-spatial mapping having a higher workload assessment than the spatial (M = 34.3% vs. M = 24.0%). This evidence supports H3.

When asked which condition was preferred, 10 participants chose spatial, 4 had no preference, and 2 preferred the non-spatial. A Chi-square test comparing the participant preference shows that the preference for spatial is statistically significant ($\chi^2(2) = 6.5$, $p < .04$), therefore supporting H4.

3.3.7. Discussion of Experiment 1

The results from Experiment 1 answer the first question stated in Section 3.2. The data indicates that the use of a layout in the interface that corresponds to the physical layout of the destination results in faster interaction and fewer errors. This result is quite straightforward and confirms the hypotheses formulated from the Dimensional Overlap model; however, this relationship had never been experimentally tested in the context of human-computer interaction, let alone in multi-display object movement scenarios.

The measures reveal that using a non-spatial technique with a spatial task could mean a delay of 29% in time and more than three times the error rate in common cross-display object movements. These results should, however, be interpreted with care: in a real scenario there is generally no constant need to transfer files one after each other, and there are several other factors that may vary (e.g., the number of displays involved). Although these extra questions require further testing of the hypotheses, it is likely that the effects revealed by this experiment will actually increase when CDOM is occasional, or if more users and displays are involved. If the scenario includes many more displays or people, the search of the correct name for the non-spatial technique will take a longer time, whereas the task with the spatial technique will remain essentially the same (except that it might, perhaps, require a bit more precision to target the right destination). If CDOM happens in the middle of other activities, users will need to refocus on establishing the correct mapping, which is likely to incur more overhead for the non-spatial mapping which, as shown by the results above, causes a larger workload.

It should also be noted that the spatial condition of Experiment 1 contains the non-spatial condition; i.e., users could have decided to ignore the spatial mapping of the application and use only the names in the circles. Observations during the experiment and the comments of the users suggest that, when a spatial mapping is available between stimuli and actions, this mapping is automatically chosen against the name mapping (probably because people will use the method that causes the least cognitive workload). This is supported by participant comments stating that "it is easier to just click the spot where you hear the person, rather than looking at the names", and that "[with the spatial condition] it is a lot more easier".

It is important to note that the layout of the application was identical for both conditions (i.e., the arrangement of the target circles was the same for spatial and non-spatial – see Figure 14); the non-spatial condition did not require such an arrangement. Instead, the non-spatial technique could have used a different kind of layout (e.g. linear from top to bottom). The choice to maintain the same layout across conditions responds to the need to make the conditions comparable, and avoid possible confounds (e.g., Fitts's Law[6] can have a large effect in targeting time, that would most likely make recorded times different for the conditions). It does not seem likely that a different layout could influence the results differently for each condition; however, the current experiment cannot rule out this possibility and further studies can be helpful to discard this possibility. The same experimental design decision also holds for Experiment 2.

Finally, in Experiment 1 and Experiment 2 (which will be discussed in Section 3.4) each display is associated to a person. It is possible that spatial and non-spatial techniques would perform differently when referring just to displays than when referring to collaborators (for example, through the *Social Facilitation* effect, which has been shown to affect arousal and workload[7] [Zajonc, 1965]). Further studies can help determine whether our results are affected

[6] Fitts's Law is discussed in more detail in Section 6.1.1.

[7] The Social Facilitation effect is the phenomenon that performance of individuals is affected by the presence (and knowledge by the performer of the presence) of other people observing the task. In general, the presence of others increases performance if the task is easy, but decreases it if the task is difficult.

by social facilitation; however, it seems reasonable to assume that the effects shown are mostly due to the dimensional overlap characteristics of the task and that, if social facilitation is an important factor, it would only contribute to increase the differences between spatial and non-spatial techniques.

3.4. Experiment 2: Non-spatial Task

Experiment 1 and 2 are based on the comparison of two basic techniques: in the non-spatial technique the only relationship between the destination and the interface are the names displayed in the circles; in the spatial technique stimuli and action are linked by the names *and* the spatial relationship (i.e., spatial and non-spatial encodings). In a non-spatial task, the Dimensional Overlap model described in Section 3.1.4.2, would predict that the spatial and non-spatial techniques would have to perform equally fast, since stimuli and response would not overlap in any of the two conditions. However, the action still takes place in a physical setting where destinations have physical locations as well as names. The question that this experiment tries to elucidate is whether the spatial layout of the environment affects the performance of the task, even if the spatial information is not necessary (or useful) for the task.

In order to answer this question the task needed to be redesigned to eliminate the relation with the physical environment of the task of Experiment 1. For this, I introduced a level of indirection in how the stimuli are presented. Instead of saying “me”, a symbol appears on screen that is associated with a destination at the beginning of the block. This task tries to reproduce a situation in which users have to distribute a document to the right person according to its type (e.g., text documents to John, spreadsheets to Peter, etc.).

3.4.1. Apparatus

The experiments took place in the same location and with the same devices than Experiment 1. The interface was similar to Experiment 1, but the homing location in the middle of the screen was also used to represent an abstract symbol out of the set shown in Figure 17.

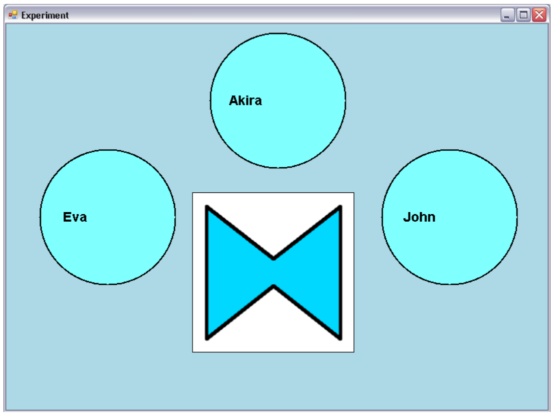

Figure 16. The interface window for Experiment 2
(this would be the spatial condition for the participant seating order of Figure 13).

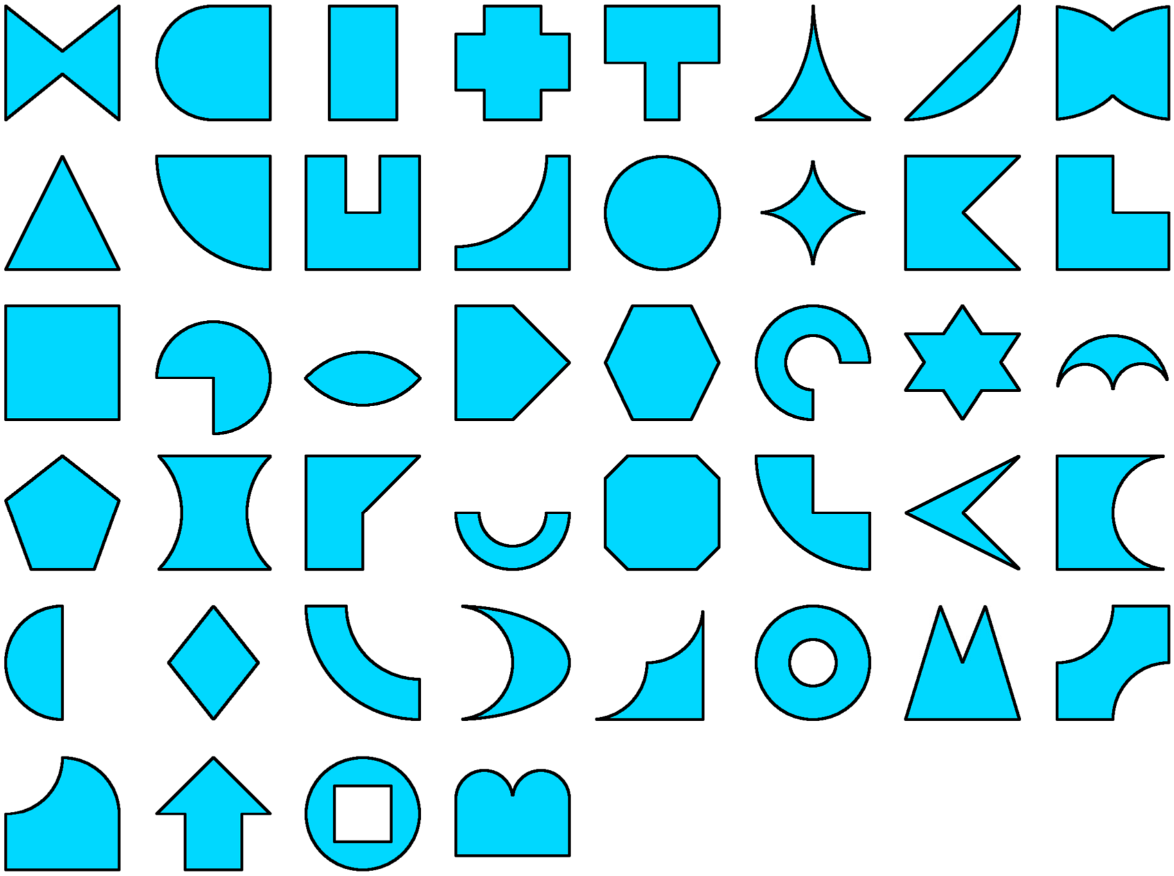

Figure 17. Symbols used to indicate the destination.

3.4.2. Participants

The same participants from Experiment 1, in the same groups, took part in Experiment 2.

3.4.3. Task

As stated before, in this task an extra level of indirection was needed to assure that the stimuli and the task would not be spatially related and to reproduce the relationship between object and destination of a real situation. This was done with the help of abstract symbols that indirectly referred to the different destinations.

At the beginning of each block the participants switched seats and were assigned a unique symbol for the whole block. Symbols were randomly selected from the set shown in Figure 17. Then all the participants showed their screen to the other participants, which displayed a large version of the symbol assigned to the owner, so that the participants could associate the right symbol to the right person. Each symbol was only used for a single person, and once a symbol was used in a block, it did not appear again in the experiment.

For each trial, a shape appeared on the screen. Users had then to remember whose shape it was, touch on the shape (homing) and then on the corresponding named circle. A participant never got to match her own shape. For each trial the system indicated if the participant had made an error or not. The homing touch was introduced to avoid variability due to the different distance of different targets to the initial position of the participant's hand.

The locations of the circles themselves were always the same (same as in Experiment 1 – Figure 13), but the names in the circles would correspond to the position of the other participants in the spatial condition (the name in the left circle would be the name of the person seating to the left, the name in the top circle would be the name of the person in front, and so on), whereas in the non-spatial condition none of the names matched the positions of the fellow participants.

The participants went through a training process before being tested in each of the conditions. The purpose of the training was to reduce the variability of the data due to learning of the task, that is, to reduce large differences between initial and final trials that arise from the familiarity of the tapping task. For training we used a variant of the task in which, instead of seating facing each other, participants sat by themselves. In the training task the names were not those of the other participants, but randomly chosen names of five letters instead. The associations between shapes and names were communicated to the participants at the beginning of the training block through a popup window.

3.4.4. Hypotheses

The experiment tested the four hypotheses of Experiment 1, and a specific one related to the learning of the mapping (H1b). The extra hypothesis was formulated after observing results from the pilot experiments that indicated that completion time stabilized somewhere between the first 10 and 20 trials. Note that the hypotheses are formulated assuming that the physical setting does interfere with the task, even if the task is not spatial.

H1: Participants take less time to tap the corresponding circle in the spatial condition

H1b: Participants take less time to learn symbol-person associations in the spatial condition

H2: Participants make fewer mistakes in the spatial condition

H3: Participants report less workload for the spatial condition

H4: Participants prefer to perform the task in the spatial condition

3.4.5. Experimental Design

Participants were first shown a demonstration of the procedure in which the experimenter explained the task while reproducing the task himself. Then, participants trained by themselves (i.e., in different locations, not facing each other) in two blocks of 75 correct trials in which random five-letter names were used instead of the names of the other participants. These names were used for both training blocks. During the training the participants did not share the same space in the room and could not see each other. The order in which the different shapes appeared did not follow any recognizable pattern. For each block the shapes were different, as were the positions of the name labels.

After training, participants sat in the same arrangement described for Experiment 1 (see Figure 13). The order in which they executed the spatial and non-spatial conditions was the same as in Experiment 1.

In each condition, each group performed a training block of 15 trials and two blocks of 75 trials (counting only correct trials) for a total of 30 training trials and 150 correct test trials. For each block the participants changed their positions relative to each other, and they were assigned new shapes. If a participant made a mistake, that trial was replaced for another at the end of the block, assuring that all participants performed exactly 75 correct trials in each block, and a variable number of wrong trials.

Once all blocks were completed by all four participants, they were asked to complete a post-task questionnaire (separate from Experiment 1). The questionnaire required the users to rate the spatial and non-spatial conditions of the group task using the standard NASA-TLX performance measures. They also had to select which condition they preferred.

We took four measures: completion time, errors, subjective workload assessments from the participants, and preference.

3.4.6. Results

An ANOVA test with participant as random factor and type of technique (spatial vs. non-spatial) as within-subjects factor revealed that completion time was less for spatial than for non-spatial mappings ($F_{1,15} = 6.88$, $p < .02$). Completion time was log-transformed in order to comply with the normality assumption of the ANOVA test. However, the difference between conditions was very small: the average completion time for spatial mappings was 780ms and for non-spatial mappings 827ms, a difference on average of 47ms (or 6%).

Inspection of the data revealed a strong learning effect which favors whichever condition was performed last (see Figure 18 and Table 2). This effect shows up as a significant interaction between order and spatial condition in a repeated-measures ANOVA ($F_{1,14} = 6.87$, $p < .02$).

Although the different orders across conditions were distributed evenly (balanced design), we cannot discard the existence of asymmetric order effects across conditions (e.g., the spatial condition might be a much better "training" for the non-spatial condition than the reverse). Although I do not believe this is likely the case, I decided to nevertheless analyze only the first condition for each participant, effectively transforming the design into a between-subjects comparison. This design type is, however, much less powerful than the previous one, and cannot show any statistically significant differences due to the large variations in individual performance and the small apparent size of the effect. Therefore, these results only provide weak support of H1.

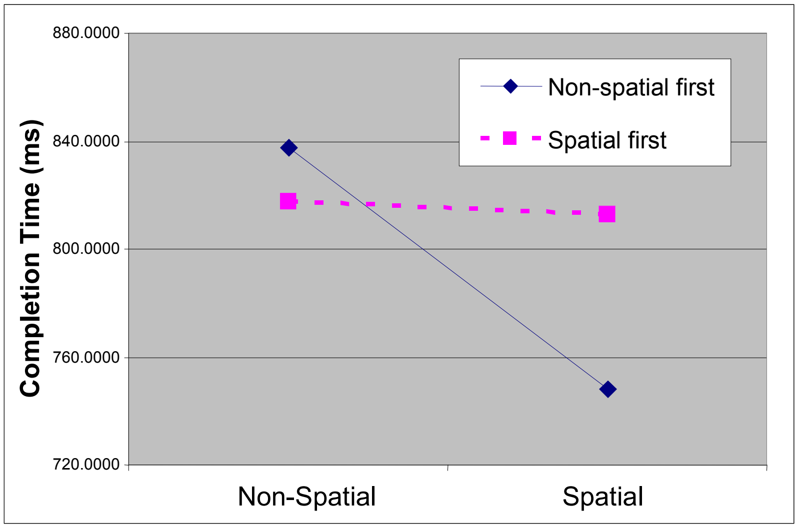

Figure 18. Order effects in the completion time data of Experiment 2.

Table 2. Average completion time from the start of the trial (measured in milliseconds) and standard error (between parentheses) for the different conditions (columns) according to the order of presentation of the conditions (rows).

	Spatial (SE)	Non-spatial (SE)
Spatial First	747.84 (7.55)	837.78 (12.22)
Non-spatial First	812.63 (10.55)	817.28 (8.9)

To test H1b, the initial trials for each position were analyzed for completion time. Figure 19 shows the first 18 trials for each position grouped in sets of six measures. The figure displays a clear learning curve for both conditions in which completion times stabilize after approximately the 12th trial. More importantly, there is a clear difference between the conditions, with the spatial condition being faster from the beginning, and the non-spatial condition eventually catching up and saturating at approximately the same level. The visual analysis is backed up by a repeated-measures ANOVA in which the condition (spatial or non-spatial) was modeled as a between-subjects variable and the trial order was the within-subjects variable. Trials were collapsed into three different groups of six trials (as in Figure 19) to reduce noise: trials 2 to 7 in the first group, trials 8 to 13 in the second group and trials 14 to 19 in the third group (the first trial was eliminated because it is subject to external factors such as setup delay). The analysis reveals a significant effect of the main condition ($F_{1,14} = 4.23$, single-sided $p < .03$), and trial group ($F_{2,28} = 14.02$, $p < .001$). In particular, the differences in average are very important in the first 12 trials (first and second trial groups), reaching a difference in means of 294ms in the first group and of 122ms in the second group. These findings indicate that learning the person-symbol relationships has a smaller cost for the spatial condition than for the non-spatial condition, and therefore H1b is supported.

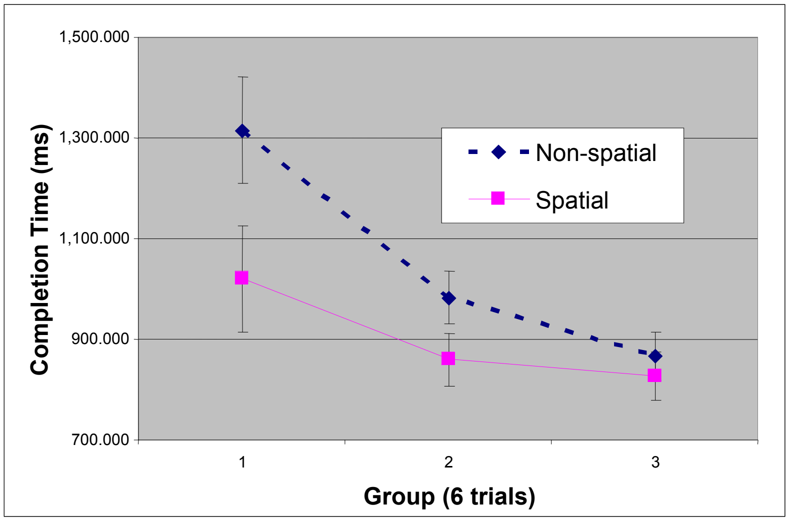

Figure 19. Average completion time for the different conditions after each position change. Blocks aggregate data from 6 contiguous trials. Error bars represent Standard Deviation.

The error data was analyzed using a repeated-measures Wilcoxon Signed Ranks test that showed no difference between the spatial and non-spatial mappings ($z = -.352$, $p < .72$). The error data was analyzed also using only the first condition, also yielding non-significant results. Therefore, the error data does not support H2.

The TLX workload assessment was analyzed for differences between the spatial and non-spatial mappings. A Wilcoxon Signed Ranks test revealed no statistically significant difference between the two conditions ($z = -1.470$, $p > .14$). It should be noted that the subjects answered this questionnaire only after doing all the trials, which means that the data reflects the subject's perceptions across all trials, not only the initial ones where the log data finds statistical differences. H3 cannot be supported by this data.

For the question of which mapping they preferred, 7 participants chose the spatial mapping, 3 chose the non-spatial, and 6 stated no preference. These differences were analyzed using a Chi-square test that reveals no significant difference ($\chi^2(2) = 1.62$, $p > .44$). This means that H4 is not supported by this data.

3.4.7. Discussion of Experiment 2

The results from Experiment 2 show little support for an overall advantage of spatial techniques for tasks where the encoding is non-spatial. This result and all the accompanying negative results of errors, subjective workload and preference are consistent with the predictions from the Dimensional Overlap model: if the encoding of the task is non-spatial, a technique which allows either a spatial or non-spatial encoding of the action will not have a better dimensional overlap than the non-spatial technique, and therefore their performance will be equivalent.

However, the data does show that associating stimuli with destinations does happen faster in the initial trials for the spatial condition (while the participant is still not very familiar with the symbol assignments), and that it takes less time to stabilize. Although the task of Experiment 2 is not strictly an stimulus-response task (it has a memory component), the results are relevant for the design of interaction techniques because the task is similar to tasks that take place in the real world (see example in Section 3.4). These results might be caused by an inherent advantage of spatial recall (although this would contradict Jones and Dumais' results [1986]), or by the possible interference of the spatial environment into the task. Anecdotal evidence from the pilot studies and my own experience in testing the software suggests that, especially at the beginning, the automatic reaction was to perform a spatially related gesture (according to the physical location of the intended display), and it was only after some practice that this natural reaction could be compensated for.

As with the previous experiment, we do not know if the results will hold for more participants or in situations where the transfers only occur occasionally. However, it is likely that the learning/memory effects will only be magnified by the increased difficulty and by having to remember while in the middle of other tasks, which would act as distractions.

Finally, the spatial arrangement of the experimental interface (Figure 16) might have played a role in the differences in the learning effects, as discussed for Experiment 1 (Section 3.3.7). The layout of the circles was the same in the two conditions (see Figure 14) to avoid the possible confound of the different targeting time differences that can be expected from different layouts. Again, new experiments will have to be run to completely rule out the possibility that the learning effect is due to the inability to see the diamond layout as a non-spatial selection mechanism. In any case, this result would not advocate for exclusively non-spatial techniques (which are, at best, equivalent to techniques that allow the two encodings), but instead just warn against the two-dimensional representation in the interface of objects and people that are also visible in the physical environment.

3.5. General Discussion and Open Questions

During the last few decades of ubiquitous computing and co-located groupware research there has been a clear emphasis on providing interactivity that is location dependent - that is, systems where interaction techniques are spatial and require the sensing of spatial relationships between users and devices (e.g., [Manesis & Avouris, 2005; Kortuem et al., 2005]). However, there is little experimental evidence to show whether or why spatial interaction is superior to non-spatial interaction.

The results from Experiment 1 and 2 provide empirical support for the use of spatial techniques in the transfer of objects across displays in MDEs. The results show that, when the task is encoded spatially, CDOM is faster, causes less perceived workload, and results in fewer errors. When the natural encoding of the task is not spatial, using a technique that has a spatial component on top of the non-spatial mechanism (naming) does not negatively affect performance; on the contrary, in some cases it might be detrimental to neglect the spatial dimension of the system in the design of interaction techniques (e.g., if associations between people and interface elements must be memorized).

Sections 3.3.7 and 3.4.7 provide detailed discussions of how tradeoffs in the design of the experiments require that we still interpret these results with caution. However, the evidence points to a substantial performance gain if spatial techniques are used and, as discussed above, we can expect differences between spatial and non-spatial to be even larger if more people or displays are present or if the CDOM takes place in the context of other activities.

It is difficult to exactly predict the impact that this design choice will have on real systems. The answer will likely depend on the particular situation and the general task for which the system is built. If the tasks are mostly spatial or transfers of objects are extremely common (e.g., in very dynamic and cooperative video games), the design of the interaction technique might be fundamental to provide a fluid interface; if CDOM is rarely used, or if most of the sub-tasks required are abstract and non-spatial, using an spatial technique might be desirable but not critical. Differences between individuals might also play a role. For example, spatial tasks are only so because a spatial encoding of the information is straightforward and natural; however, some individuals might naturally use different perceptual mechanisms, or might consciously choose to encode the destination in a different way (e.g., users might decide to consciously ignore spatial information and focus instead on names). Although it might be useful to consider these cases and new typologies in the future (especially for groups such as persons with disabilities, children, or older adults), I believe that the results shown can be generalized across most people.

As I pointed out above there are still many open questions that require further research; in particular the effects on performance of the number of possible destinations and of the context in which CDOM takes place. As multi-display interaction becomes commonplace it will become important to refine our knowledge and determine what level of correspondence between the interface and the physical space is required to show advantage.

Finally, it should be noted that the categorization of techniques and tasks into spatial and non-spatial might be general, but is certainly not the only possibility. The rest of this dissertation is focused on the further categorization of spatial techniques; however, although most existing techniques are spatial, there might be opportunities for the design and study of novel non-spatial techniques that are useful for specific scenarios.

3.6. Implications for the Design of CDOM Interaction Techniques

The results and discussion in this chapter support the design of spatial interfaces. Systems that are aware of the position in space of all its displays and users might, however, be expensive to implement. Designers of co-located MDEs and interaction techniques should carefully consider whether the possible performance advantages of spatial arrangements justify the implementation of expensive sensing mechanisms. The cost of position sensing is, however, constantly declining and centimeter-resolution position sensing systems are likely to be affordable in small devices in a few years. Alternatively, in some situations it is possible to build techniques that ask users to specify the spatial relationships themselves; this design alternative is particularly interesting if the MDE configuration remains relatively stable.

3.7. Conclusion

In this chapter I have identified that the way in which CDOM interaction techniques refer to displays is an important design factor that affects human performance. I distinguish two main categories of techniques: spatial and non-spatial. With the help of dimensional overlap theory and the results from other experiments I hypothesized that performance in CDOM techniques depends on how the technique matches the encoding that is required by the task (e.g., spatial techniques will improve performance for tasks that are usually encoded spatially).

Two experiments tested two aspects of the spatiality of techniques and the tasks that are relevant for CDOM scenarios. Experiment 1 showed that completing cross-display object movements with a non-spatial technique took 30% longer than with a spatial technique in a spatial task. Accuracy was also superior for the spatial technique (an average of only one-third the non-spatial technique), and participants felt that using the spatial technique required less workload.

Experiment 2 could not find general differences in performance for a non-spatial task. However, subtle effects were found that suggest that, at the beginning, it is easier to establish relationships between gestures and symbols when the gestures correspond spatially to the symbolized destinations.

The results from Experiments 1 and 2 support the implementation of CDOM interaction techniques that are based on the system's spatial awareness of the physical space. Further chapters explore some of the more subtle differences in spatial arrangements; for example, Chapter 4 and 5 look at the correspondence between the spatial arrangement of the displays and the control actions of the user, and Chapter 6 explores the different factors in the execution of CDOM for spatial interaction techniques.

CHAPTER 4: MOVEMENT PLANNING - FINDING THE RIGHT CONTROL MOVEMENT FOR THE RIGHT DISPLAY

In spatial terms, an MDE is defined by its display configuration. The display configuration depends on both the physical arrangement of the MDE (the positions and physical properties of its displays), and the input model of its interaction techniques (the correspondence between control movements and the movement of the object between displays). The relationship between the physical arrangement and the input model (which I henceforth call the mapping) can be designed in many different ways for a given MDE that has a particular physical arrangement. This chapter analyzes how the different mappings of CDOM techniques affect the movement planning process.

To illustrate the large number of possible mappings we turn to a simple example. Consider the environment shown in Figure 20. If this environment is controlled through a traditional mouse-cursor interface, the mapping matches motions of the mouse over the table with the movement of the cursor across the horizontally arranged set of displays. Because the input is indirect, it is possible to create different mappings between the movement of the mouse and the movement of the cursor: a traditional mapping will associate left-right movements of the mouse with horizontal movement of the cursor, and forward-backward movements with up-down movement of the cursor. Even at this level it is possible to create different (although possibly not desirable) mappings (e.g., switch the relationship between forward-backward and up-down, etc.). However, it is possible to create more mappings by using different kinds of input. For example, we can consider laser pointer input (e.g., [Myers et al., 2002]) where the mapping is established geometrically as the intersection of the line coming out of the input device with the display (see Figure 21).

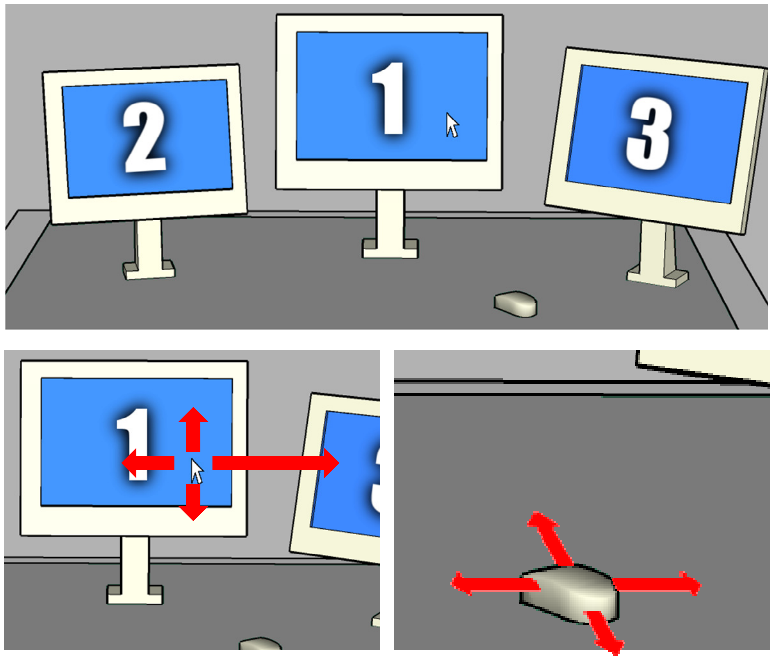

Figure 20. The common mapping of a cursor-controlled mouse in a simple MDE: forward-backward movements of the mouse cause vertical movements of the cursor, left-right movements of the mouse cause horizontal motion of the cursor.

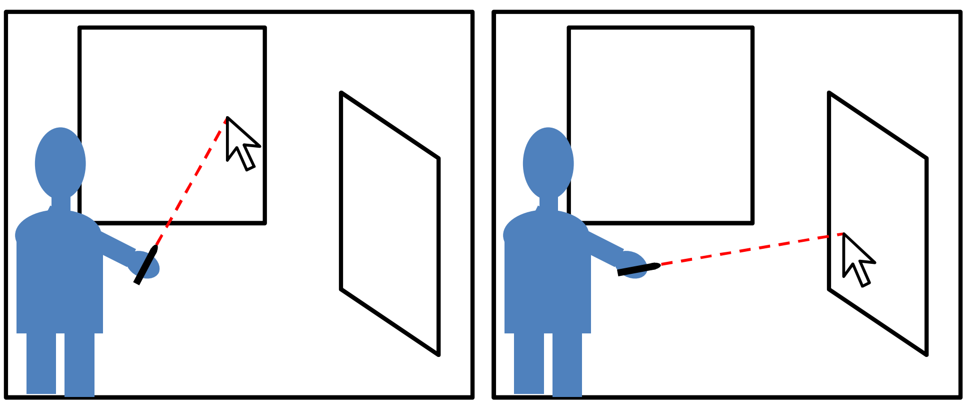

Figure 21. A generic laser pointer technique for cursor control.

A rich set of possible spatial techniques emerges from the different ways in which the mapping between the physical space and the virtual space is designed; the question now arises of how the different mappings will affect human performance when trying to move an object from one display to another.

4.1. Planar, Perspective and Literal

The discriminating characteristic that I use at this level to differentiate CDOM techniques is the type of input model underlying the technique. I distinguish between three basic models: planar, perspective and literal.

4.1.1. Planar

Planar input models are the most common in current multi-display environments. In planar interaction techniques the input space of all displays is stitched together in a flat surface. Current operating systems configure multi-display interaction in this way (see Figure 22, [Apple Corporation, 2008; Microsoft Corporation, 2008]) because this kind of input model allows a simple one-to-one mapping of the two-dimensional movement of common input devices (mice, touchpads, trackballs, pen digitizers) into the available display space. This kind of input model seems adequate for simple MDEs like current multi-monitor PCs, where displays are usually arranged in simple physical configurations (usually a vertical plane that covers a small fraction of the visual angle and is perpendicular to the user).

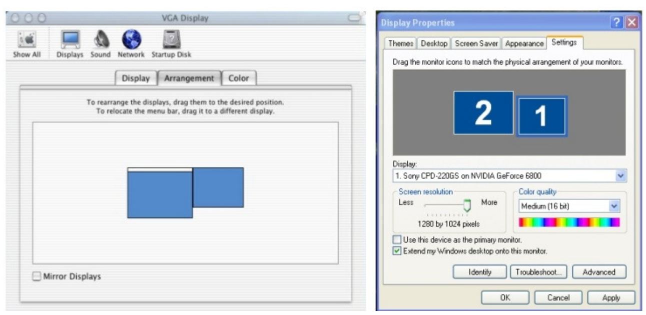

Figure 22. Multi-monitor mapping dialog of two operating systems.

Other planar interaction techniques have emerged to represent more complex MDE configurations through different strategies. For example, Mouse Ether [Baudisch et al., 2004] is a refinement of the common multi-monitor technique that is designed to improve the match between input space and physical configuration. Mouse Ether can represent the physical configuration more faithfully because it accounts for the physical space in between displays and for the differences in resolution between displays.

PointRight [Johanson et al., 2002] is another extension of planar stitching[8], where displays are no longer in the same plane, but instead cover different surfaces of a room. In PointRight the cursor behaves similarly to a regular multi-monitor setting, but the pairings of transition edges between displays are chosen so that users can easily understand where the cursor has to go out from one display to appear in another one (see Figure 23).

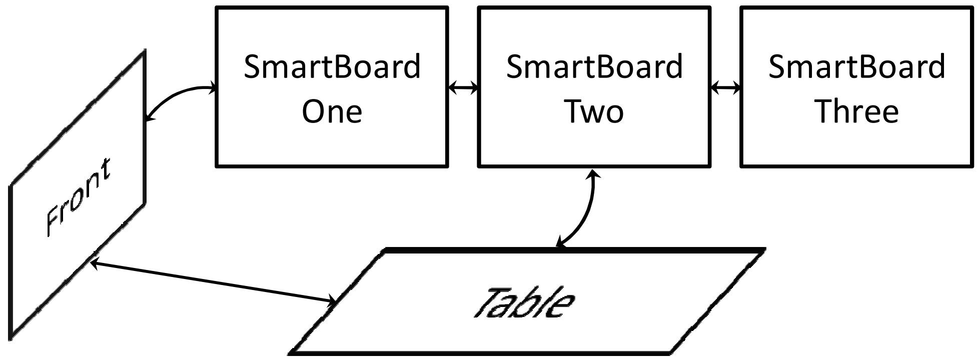

Figure 23. The mapping used by PointRight (reproduced with permission from [Johanson et al. 2002]).

Swordfish [Ha et al., 2006, 2006b] works under the same basic planar assumption, but the edges of the displays are connected to each other in different ways that can be arbitrarily chosen by the user, i.e., the geometry of the input model is configurable in real time.

ARIS [Biehl and Bailey, 2004], allows the windows of different applications in a multi-display room to be moved across displays through direct interaction with a miniature representation of the room display space. This representation flattens the volume of the room into a map (see Figure 24).

Many other techniques assume a flat mapping of displays. For example, Drag-and-Pop [Baudisch et al., 2003] creates proxy objects that allow access to distant objects that are in other displays from the area where the user is interacting. The distant objects and their respective proxies are connected through a rubber band graphic that keeps the visual connection (see Figure 25). Because proxies are only created for objects that are located in the direction of the gesture, the system must have a sense of where displays are located with respect to each other in the plane of interaction. Other related techniques such as Vacuum [Bezerianos and Balakrishnan,

[8] In this dissertation, the word "stitching" is used to refer to a way of configuring the control spaces of displays within the same plane (even though the physical displays might not be necessarily aligned in this way) so that the space can be traveled in a two-dimensional way. I indicate explicitly when the term is used to refer to a different concept.

2005], Frisbee [Khan et al., 2004], Push-and-throw and Pantograph [Hascöet, 2003], Flick, Superflick [Reetz et al., 2006], and MDE versions of the Radar [Nacenta, 2007] similarly use an underlying planar input model, regardless of the input device that they are designed for (mice, touch, pen, etc.). More sophisticated versions of the same idea use images from a camera [Chiu et al., 2003], or from a synthetic model of the environment [Massó et al., 2006] as planar models.

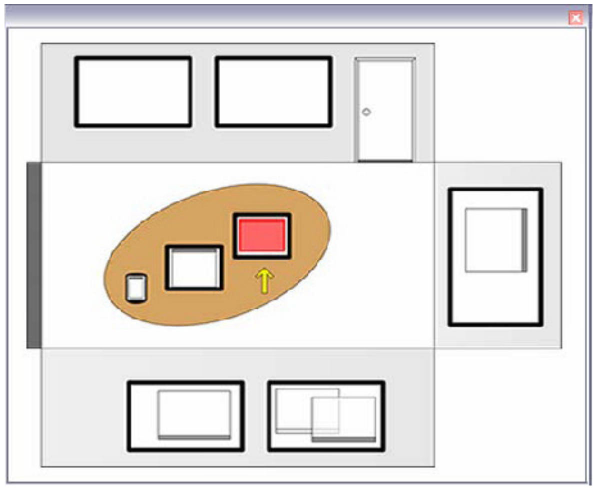

Figure 24. ARIS's flattened representation of a MDE (reproduced with permission from [Biehl and Bailey, 2004])

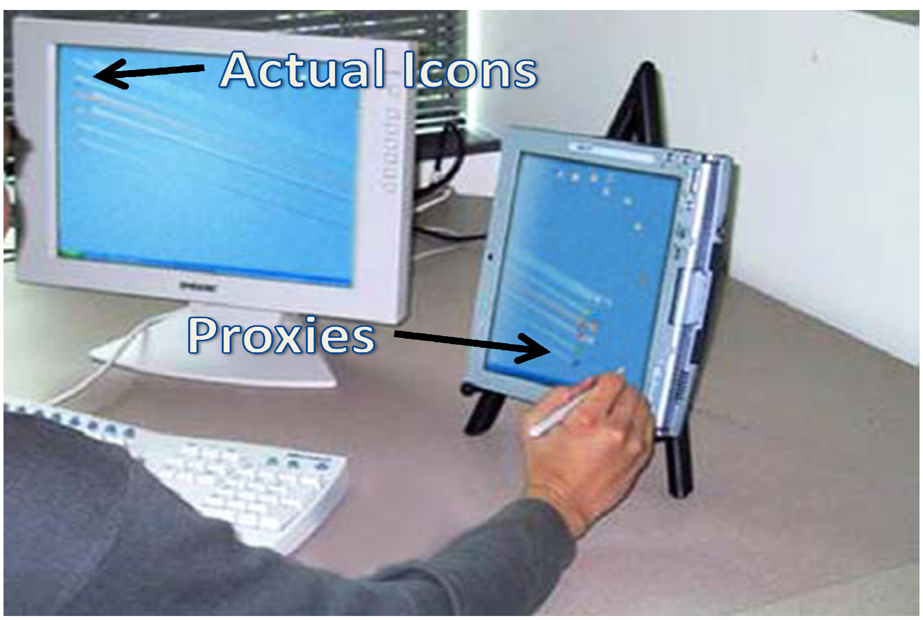

Figure 25 . In Drag-and-Pick [Baudisch, 2003], the icons that are in the other display are made accessible on the tablet through proxies. The input model of this technique reproduces the planar arrangement of the displays (image used with permission).

For other techniques such as the button-based Multi-Monitor Mouse [Benko and Feiner, 2005, 2007] and the Spatially Arranged Folders (see Figure 10), the underlying planar input model is simpler because they rely on discrete destination selection: the Multi-Monitor Mouse moves the cursor to a display when the key assigned to that display is pressed (with the keys arranged in the same spatial order than the monitors), and the Spatially Arranged Folders act as Wormholes (visual icons that, when dropping an object on top, warp it to a particular display). In these two examples, the input model is still planar because the keys and the icons are arranged on flat surfaces.

Planar input models are very common because they are easy to implement and to understand, at least for simple environments. However, planar models have important drawbacks when the MDE is complex. Consider, for example, configuring an office MDE with many displays of different sizes, resolutions, positions and orientations (like the one in Figure 4.D, page 5) through a planar model. The configuration process becomes complex, but most importantly, the result would be difficult to navigate (i.e., figuring out which edge of which display leads to which other display becomes difficult). For many MDEs it might not even be possible to create a planar mapping of the space that is consistent for all users or for all possible positions of the user; for example, the simplified environment of Figure 26 shows how the correct planar mapping varies depending on the position of the user.

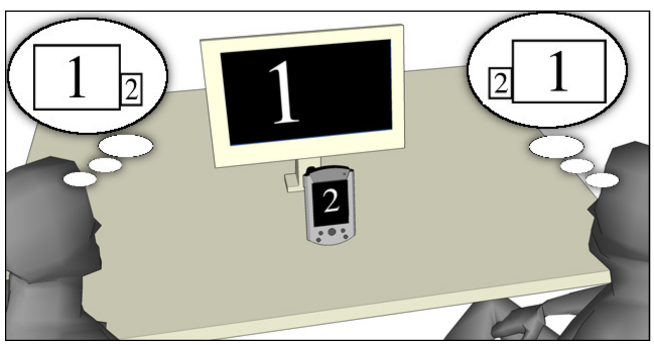

Figure 26. The correct planar mapping depends on the position of the user.

The problems mentioned above arise because static 2D representations of space are inherently incapable of faithfully represent a dynamic 3D environment.

4.1.2. Perspective

Perspective interaction techniques are those where the perspective of the user on the MDE's physical configuration somehow affects the mapping. Here the term *perspective* is used loosely to denote the way in which the user perceives the environment.

The main assumption in this group of techniques is that, since a static 2D mapping of a complex 3D physical display configuration is cumbersome to operate, a mapping that is based on how the environment is perceived by the user will make CDOM operation easier. In other words, deciding which gesture to make will become easier for the user if the input model matches what she sees, and not some arbitrary static representation of the space.

The most common perspective interaction technique is Laser Pointing (also known as Remote Pointing, Laser Beam, etc.) Note that with this technique (Figure 21, page 46) the mapping varies depending on where the laser pointing device is, which, in turn, is dependent on the position of the user. To illustrate this, consider a user pointing at a display from a close distance and the same person pointing at the same display but from a more distant position. In the former scenario, the rotation of the device that covers the width of the display is wider than in the latter scenario because it is much closer. Variations of laser pointer techniques can be found in [Myers et al., 2001; Oh and Stuerzlinger, 2002; Voida et al., 2005; Bolt, 1980; Parker et al., 2005].

An interesting variation of the perspective idea is Shadow Reaching [Shoemaker et al., 2007], where the direction of the pointing is not related to the orientation of any device, but is determined instead by two points, a source of light, and the physical embodiment of the user (see Figure 27).

Although laser pointer techniques are perspective techniques because the mapping depends on the position of the user, there is still a discrepancy between this mapping and the user's perception due to the different positions of the pointing device and the eyes of the user (parallax). To solve this discrepancy, I designed Perspective Cursor, which dynamically creates a mapping that matches the visual perception of the environment to the movement of a mouse cursor by using eye position tracking. The technique results in a seamless transition of the cursor from one display to another, preserving the visual continuity of the movement. This technique is explained in detail in Section 5.3.3.3 of this dissertation.

Figure 27. Shadow Reaching, manipulating objects through shadows (used with permission from [Shoemaker et al., 2007]).

Some perspective techniques have been proposed that use head orientation to determine which display to activate for input. For example, the Multi-Monitor Mouse variant with head tracking technique [Benko and Feiner, 2005], the multi-monitor technique presented in [Ashdown et al., 2005] and Look-to-Talk [Oh et al., 2002].

It is also conceivable to use eye-tracking technology instead of head tracking technology as done by MAGIC [Zhai et al., 1999], and Attentive User Interfaces [Shell et al., 2003, Dickie et al., 2006]. Although many of these techniques have not been proposed as CDOM techniques, the adaptation to MDEs would be trivial.

Finally, Augmented Reality researchers have also experimented with the idea of using tracked head-mounted displays to facilitate or enhance interaction with multiple displays as in the InVision [Slay et al., 2003, 2004] and the EMMIE [Butz et al., 1999] systems. A wide variety of practical problems prevent these techniques from being feasible in most current scenarios.

In many cases, perspective techniques can be described in terms of image-plane interaction [Hill and Johnson, 2008]. However, all perspective techniques share the common characteristic that they somehow adapt the mapping to the user, often dynamically, so that perception and mapping coincide. Naturally, this requires specific sensing or input devices that acquire knowledge about the user's perspective on the system. In general, this data can be provided by commercially available devices (e.g., eye-trackers [Tobii Technology, 2009; IT University of Copenhagen, 2009] and 3D position trackers [Intersense Inc., 2009; Polhemus, 2009; Vicon 2009]), but these are currently still expensive, and this might prevent the popularization of these techniques in the near future.

It is also important to consider that, although the adaptation to a user perspective can be beneficial for the user, it can also negatively affect collaboration when working in a group. For example, some perspective techniques might make it difficult for other users to predict or stay aware of the user's actions. This tradeoff is common in groupware system design [Gutwin and Greenberg, 1998]; this issue is further discussed Section 5.4, page 100.

4.1.3. Literal

The last class of CDOM interaction techniques according to the input configuration is literal techniques. Literal techniques are those where the input model coincides with the physical configuration. For example, Rekimoto's Pick-and-Drop technique [Rekimoto, 1997] (pictured in Figure 9, page 21) is a literal technique because the input model of the system (the input surface of all involved displays) coincides with the physical configuration (the displays themselves). In fact, literal techniques are just an extension of direct input [Buxton, 1983] to multiple devices and displays.

The idea of linking physical elements with their virtual effects according to their physical contact is not new; at least two current paradigms in HCI base the interaction model on it. Tangible computing in all its different forms [Ishii and Ullmer, 1996; Fjeld et al., 1999] links physical objects with the information they help manipulate, usually at the same location where display and physical prop make contact. Pen-based computing [Hinckley et al., 2009; Agarawala and Balakrishnan, 2006] relies on the fundamental assumption that direct contact of the pen with artifacts on the screen (e.g., icons, text, or digital ink) provides advantages for interaction.

Tangible computing and pen computing are not generally focused on the problem of cross-display object movement, however, there are multiple instances where research has addressed CDOM explicitly. In Smart-Its-Friends [Holmquist et al., 2001] shaking together two devices establishes connections between them allowing object transfer[9]. Hinckley's Synchronized

[9] Although the original paper does not directly address the problem of object transfer (shaking objects together creates connections between devices, but does not transfer any objects), it is

Gestures [Hinckley, 2003] work in a similar way to Smart-Its-Friends, but the connections are established through bumping tablet PCs into each other. In BlueTable [Wilson and Sarin, 2007], devices that are placed on top of an interactive tabletop system (in this case, Microsoft's Surface [Microsoft, 2009]) allow a common space to share pictures and documents between the two displays. Hinckley's *Stitching*[10] [Hinckley et al., 2006] allows for the sharing of documents and pictures from one display to another by detecting the relationship between pen strokes that are performed across the two displays (see Figure 28).

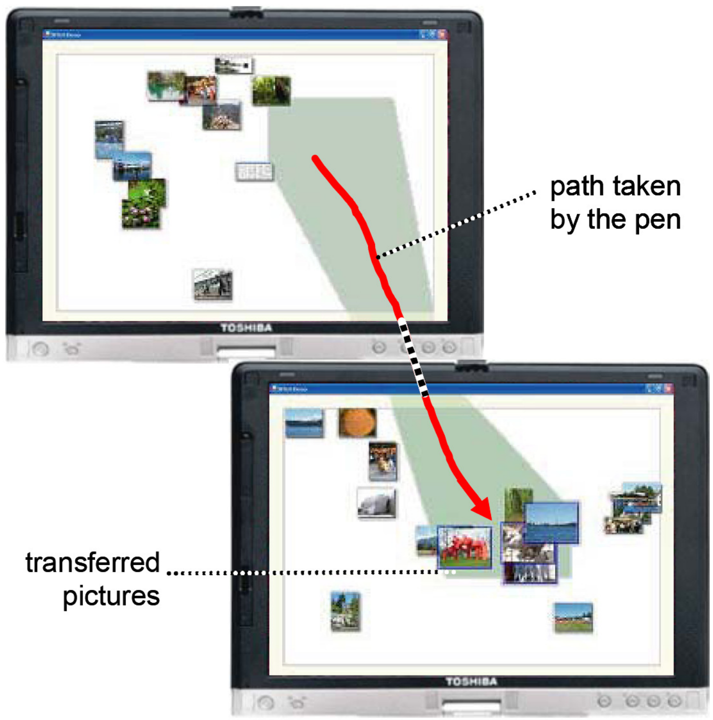

Figure 28 . In Stitching [Hinckley et al., 2006] a pen traced over two displays is used to transfer pictures from one pen-based PC to another.

In Passages [Streitz et al., 1999], physical objects can be used to transfer information between displays (e.g., a large vertical display and tabletop display) as long as they are equipped with a bridge device (a device that can uniquely identify a physical object through the sensing of some variable – in their initial implementation they used weight).

---

trivial to extend this idea to object transfer. For example, the user selects an object in one of the two devices (e.g., PDAs), then shakes the two devices together, and the object is automatically transferred to the other display.

[10] Here Stitching refers to Hinckley and colleagues' technique, not to the more general technique that assigns transitions from the border of one display to another (see also footnote 8, page 47).

Other researchers have also used the sensing of physical proximity or physical connection between displays to establish planar-type links between the displays that allow for the transfer of objects. In u-Textures [Kohtake et al., 2005], special panels of flat displays would arrange themselves into composite displays when physically connected. A previous, but similar idea was explored in ConnectTables [Tandler et al. 2001], where several small physical desks could detect the proximity of other ConnectTables and automatically configure them as if they were a shared space.

World-in-miniature or radar techniques that are used with a direct input device (e.g., a pen) can also be considered literal techniques as long as the miniature can be manipulated independently of the real object it represents. In these circumstances the world-in-miniature proxies effectively turn multi-display interaction into a simple single-display direct contact action, which can leverage the advantages of direct input (see the discussion in [Nacenta et al., 2009]), but with the added advantage that the space to be covered by the user is reduced (instead of having to directly touch the destination, the whole interaction happens within the much smaller world-in-miniature). However, this introduces the problem of matching the miniature representation with the real object in space: it might be difficult to identify in the miniature which one is the object that has to be moved across displays, and it might be difficult to match a representation of a display in a small diagram with its physical counterpart.

Literal techniques are appealing because they resemble how people manipulate the real world. However, this comes at the cost of *power* (as defined at the beginning of Chapter 2, page 8); literal interaction techniques are restricted to the range that is comfortably reachable by the user, whereas perspective and planar techniques are not necessarily bound by this restriction. The following sections provide an overview of the issues and trade-offs that differentiate the three groups of techniques, and review existing work relevant to the topic.

Now that the main three groups of techniques at this level have been presented, we look at existing knowledge that can help us understand the principles that will determine the techniques' performance.

4.2. Optimal Mappings

It is a well-known HCI design principle that the input of a system should be designed to match what is to be controlled; Norman calls this "getting the mapping right" or designing “natural mappings” [Norman, 2002] and illustrates it with the example of the controls of a stove (Figure 29), Britton and colleagues [1978] call it “kinesthetic correspondence”, and Jacob and colleagues apply a similar philosophy to the design of general interaction techniques [Jacob et al., 1994]. In cross-display object movement techniques, this principle translates into mapping the input so that it matches the arrangement of displays in the physical space.

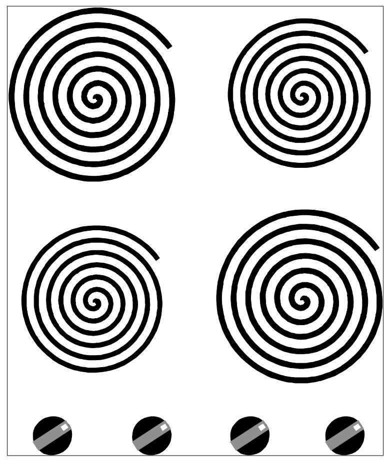

Figure 29. The classic problem of the stove. Which knob corresponds to which burner?

As pointed out above, if the MDE is very simple and regular in its arrangement (e.g., a large composite wall display), planar techniques are easy to match to the space. However, when the MDE’s physical configuration is more irregular, the match becomes more difficult. Irregular configurations are bound to happen in real environments when the MDE includes heterogeneous displays (e.g., displays of many sizes), when some displays are mobile, or when the location of the displays has to be adapted to the existing physical space instead of the other way around. Figure 30 shows an example of a relatively simple display arrangement that would be difficult to map well with a planar technique.

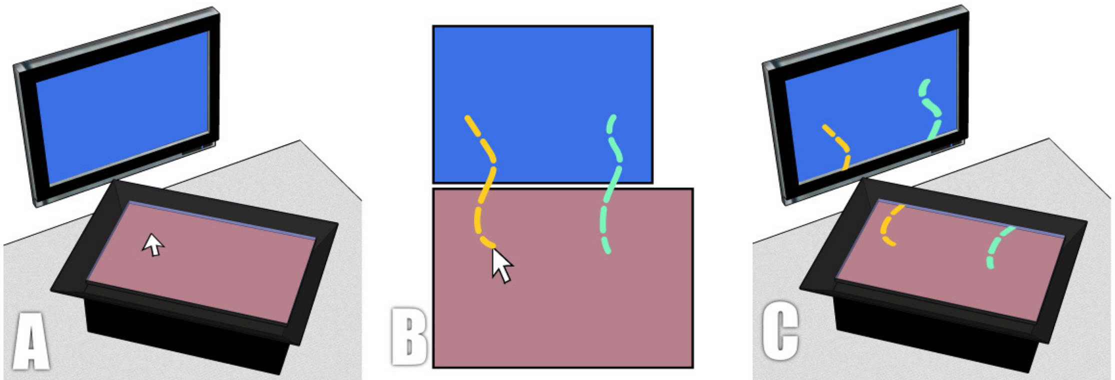

Figure 30. Two-display MDE. A) Physical arrangement, B) input model with two mouse trajectories, c) the mouse trajectories become confusing in the physical environment.

To answer the question of how the quality of the match will affect performance we have to turn again to Cognitive Psychology.

4.3. Stimulus-Response Compatibility and Dimensional Overlap

In Chapter 3 the Dimensional Overlap (DO) model was brought to attention as a way to explain why certain ways of referring to displays would be more efficient than others depending on how the representations of the displays exist in the minds of the users (see Section 3.1.4). At that level of analysis, the focus was on the type or “language” of representation of the displays: spatial or non-spatial. This chapter is focused only on spatial representations, but the DO model and its predecessor, Stimulus-Response Compatibility, are still highly relevant.

Consider Figure 29, Figure 12 and Figure 1, which are summarized in Figure 31. There is a strong commonality between the three examples. In Norman’s stove example (left), users need to decide which knob to operate to activate one or more of the four stove elements. In the typical stimulus-response task (center), the participant has to press the correct button when one of the light bulbs is turned on in the experiment. In the use of a spatial CDOM interaction technique the user will have to choose a spatial gesture that brings the object to the right display. In essence, these three tasks are equivalent; elements, light bulbs and displays provide the stimuli (or goal) of the action, and knobs, buttons and the CDOM interaction techniques establish the available responses. By learning about the results of the well-studied S-RC phenomena we can better explain why some CDOM technique designs outperform others.

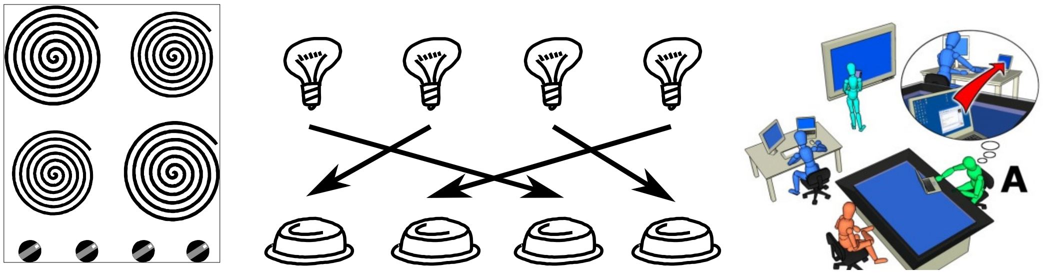

Figure 31. Analogy between stimulus-response tasks (left and center) and the CDOM task (right).

Further stretching the analogy, we can relate literal techniques to a stove where the elements are activated when directly touched, or an S-RC experiment where the light bulbs have to be directly touched when they light up. Planar and perspective techniques would correspond to setups where the relationship between the arrangement of the knobs and buttons has different correspondences with the arrangement of elements and bulbs.

As already stated in Section 3.1.4.1, mappings that are compatible will result in faster reactions and fewer errors [Fitts and Deininger, 1954, Fitts and Seeger, 1953, Chua et al., 2003, Proctor and Reeve, 1990, Proctor and Vu, 2003]. Moreover, this effect cannot be eliminated through training; after thousands of trials compatible mappings will still result in faster responses.

Applied to the context of MDEs, this effect seems to imply that if the configuration of a simple multi-monitor system is not compatible (Figure 32), cross-display movements of the cursor will be negatively affected, even after weeks of use.

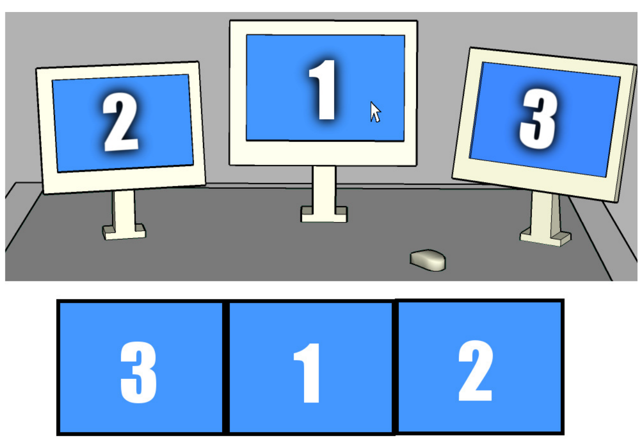

Figure 32. An example of an incompatible mapping for a multi-monitor system. Notice that the input model (bottom) does not match the physical configuration.

The discussion above suggests that we aim at compatible designs for CDOM interaction techniques. However, sometimes it is not clear which mapping is the compatible one. Although it is fairly clear which mappings would be compatible in the examples from Figure 12 (page 26) and Figure 32, it is not so clear which of the mappings from Figure 33 is compatible for the stove example.

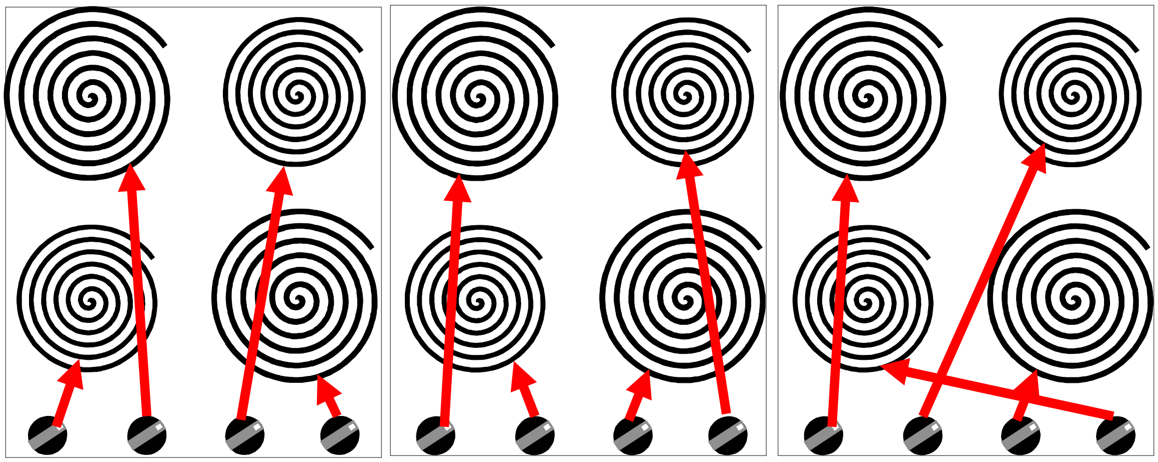

Figure 33. Three of the possible mappings between elements and knobs for Norman's stove example [Norman, 2002].

The reason is that in the stove example, elements and knobs are laid out in different dimensions (elements are in a 2D grid, knobs are in a line), whereas in the bulbs-buttons example the configuration of stimuli and possible responses are the same. Here is where the Dimensional Overlap model becomes useful; as a generalization of Stimulus-Response Compatibility, the Dimensional Overlap model postulates that S-R sets can only be compatible or incompatible if they have dimensional overlap. The bulbs-buttons example has an almost perfect dimensional overlap because the physical alignment of stimuli and response is identical. To achieve a similar effect in the stove, we can realign the knobs in two dimensions (see Figure 34) which makes the compatible mapping obvious (a solution which was already suggested by Norman and that has been adopted in the design of some modern stoves).

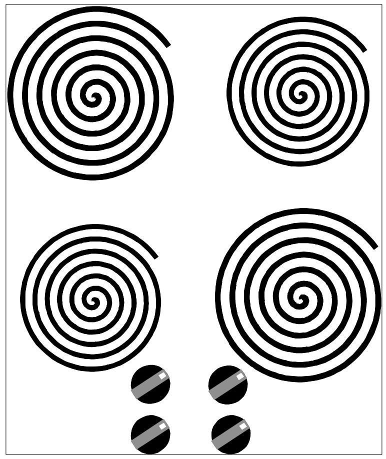

Figure 34. Realigned knobs make the compatible mapping obvious.

4.3.1. Applicability of S-RC and DO to CDOM Interaction Techniques

In CDOM tasks we usually have a physical environment composed of a number of displays that could be arranged in different ways. Some configurations are linear (1D – e.g., Figure 33, page 59), some other are flat (2D – e.g, Figure 4.B, page 5), and the most general layouts are complex and three-dimensional (e.g., Figure 4.C and D, page 5; and Figure 30, page 57). The DO model predicts that any configuration will be appropriately represented by literal interaction techniques because the overlap is, by definition, perfect. The model also predicts that planar interaction techniques will be appropriate to access planar display spaces, but will cause problems when the physical configuration is not planar – i.e., the DO model provides an explanation of why Figure 2, bottom right, (page 3) might be problematic.

Perspective techniques usually allow the control of two degrees of freedom which would suggest, at least at first sight, that they are not adequate to represent three-dimensional environments[11]. However, we should consider that visual human perception of the environment is mostly two-dimensional in that the whole environment is projected into the retina, a two-dimensional surface[12]. Perspective techniques can therefore achieve a good dimensional overlap

[11] Although laser pointer techniques allow the manipulation of five degrees of freedom (three axes for location, pitch, and yaw – roll is usually meaningless for a laser dot or for a cursor), the location is rarely used to specify destination. Instead, the pointing device is located wherever it is comfortable, and it is the pitch and yaw that are used to determine destination (2DOF). When location is used to specify points in the display, laser pointer techniques approximate literal techniques (see [Parker et al., 2005]).

[12] Technically, the environment is projected into two retinas, which allows for the visual detection of depth through the slight differences in image from the left eye to the right eye.

by mapping the input to the perceived space. Disparities between techniques in the quality of the overlap will appear due to parallax (as I explained above in relation with laser pointing techniques and shadow reaching, Section 4.1.2, page 51), or due to problems in the calibration of the sensors.

4.4. Discrete and Continuous Interaction

Most of the examples of spatial CDOM techniques presented in the previous discussion assume discrete interaction; touching a button or activating a control happens in a punctuated, isolated manner. Many spatial object movement interaction techniques are, however, continuous in nature. For example, laser pointer techniques are continuous because the movement of the pointing device continuously affects the position of the cursor (or laser spot), and the physical device can be moved fast, but the laser cannot be made to warp from one position to the next. Other techniques incorporate both continuous and discrete interaction. For example, the traditional multi-monitor capabilities of current operating systems (discussed in Section 4.1.1, page 47), are continuous in that the user continuously moves the cursor in the direction of the destination display, but discrete in that the actual transition of the cursor between displays happens only in a small fraction of a second the exact moment that the boundary of the original display is crossed. In contrast, other techniques such as the Multi-Monitor Mouse (key-press variant) are discrete in that the object is instantly warped, without moving in a continuous fashion.

Chapter 6 will discuss in detail the issues of continuous and discrete interaction as they refer to the feedback loop. However, it is important to mention at this point that there are no evident reasons why the application of the Dimensional Overlap model and Stimulus-Response Compatibility should not be valid for all spatial interaction techniques, regardless of their characteristics in terms of continuity.

However, the disparity of information between the two retinal images is small due to the small distance between eyes and, unless objects are very close to the viewer, a spherical representation (2D) of the surrounding space is still a good approximation of the information perceived by the viewer [Card et al., 1999].

4.5. Cross-display and Non-cross-display Factors

Section 4.1 contains a survey of spatial techniques of cross-display object movement that describes many techniques that are based on different types of input and representations of the physical space. It is important to highlight that the focus of this dissertation is to investigate characteristics of techniques that will make a difference in terms of multi-display interaction. The study of all the characteristics that can affect performance, even if constrained to pointing, is a very large collaborative endeavor that falls out of scale and scope of this dissertation (for foundational papers in this area see [Buxton, 1983; Card et al., 1978; Card et al., 1991]).

Focusing only on characteristics that relate to multi-display interaction has two main consequences: first, it makes the topic manageable so that we abstract and study issues that are important for multi-display interaction; second, it limits the share of performance variability that can be explained within this framework. This means that the different performances of different techniques will not be explained exclusively by the characteristics or categorizations that I deal with in this and the rest of the chapters, but certain techniques might be more efficient just because they use a different input device (e.g., a mouse is generally much faster than a trackball), or because the input is of a particular kind (direct input is faster than indirect input in certain conditions), or because some other characteristic that is independent of the number of displays that are being used.

The experiments contained in the following chapters were designed to investigate multi-display characteristics such as the mapping between input model and physical space. However, it is very difficult to compare real techniques and keep other (non-MDE) characteristics equal across all conditions because sometimes the kind of input or the particular characteristic is intrinsic to a technique. Therefore, some of the experiments (especially Experiment 3) will include conditions that vary across some of these other variables. In these cases the data obtained from the experiments will reflect the effects of some non-MDE variables as well as the main issues of interest in this dissertation. For such cases, and until we have a much larger body of work, we have to trust a careful look at the data that goes beyond the statistical analysis to be able to extract recommendations for MDE design.

4.6. Workspace Awareness in CDOM

The focus of this dissertation is on performance variables (speed and accuracy of interaction); however, there are other important measures that can be crucial in the design of CDOM techniques, such as *workspace awareness* [Gutwin and Greenberg, 1996, 2002]. This and the following chapter touch on it in their respective discussions. This subsection serves as a brief background on workspace awareness.

Workspace awareness is the up-to-the-moment understanding of another person's interaction with the shared workspace [Gutwin and Greenberg, 2002]. Workspace awareness is important in collaborative scenarios because collaborators often need information about what others are doing in order to effectively coordinate and work together. For example, it is important to know what someone else is doing if we do not want to repeat the same work, or undo some of the work done by others.

Workspace awareness and awareness in general have been an important tool for the study of distributed systems because awareness is particularly difficult to maintain when collaborators are not present in the same space, and lack of this awareness is one of the causes of poor collaboration support in most of the remote collaboration tools.

Awareness in co-located collaboration scenarios has mostly received attention as inspiration or baseline comparison for distributed systems. For example, in [Tang, 1991; Heath and Luff, 1992], co-located scenarios were studied to better understand collaboration, which would, in turn, help design better distributed systems. The study of how to design systems to support workspace awareness in co-located collaboration has only received attention when it has become apparent that workspace awareness is not completely automatic in co-located scenarios, especially if new interaction techniques are used that create some kind of disconnect between where the actors are and where their actions are taking place.

In [Nacenta et al., 2007; Pinelle et al., 2008] we showed that different interaction techniques can affect the degree of coordination and the number of conflicts that arise in tabletop collaboration. My research in this area (together with a number of collaborators) is not included in this dissertation because it is not centered on performance, and because it concerns mostly single-display groupware (although most of the findings are probably also applicable to MDEs). For more information see [Nacenta et al., 2007; Pinelle et al., 2008; Pinelle et al., 2009], or see related work on co-located collaborative behavior by others [Scott et al., 2004; Morris et al., 2006; Tang et al., 2006].

In distributed groupware as in co-located collaboration it is common to find tradeoffs between favoring the individual, for example, in terms of performance or power, and favoring the group, for example by increasing workspace awareness [Gutwin and Greenberg, 1998; Nacenta et al., 2007]. The findings of those studies have not been extrapolated to MDEs yet, but there is no reason to believe that most of the phenomena observed will not appear as well (if not become exacerbated) in MDEs. Therefore it is important for designers of MDEs to take into account the social and collaborative aspects of interaction techniques as well as their individual performance aspects.

CHAPTER 5: MOVEMENT PLANNING - FINDING THE RIGHT CONTROL MOVEMENT FOR THE RIGHT DISPLAY - EXPERIMENTS

5.1. Research Questions

The design of the experiments in this chapter is focused on finding quantitative performance differences between techniques of the planar, perspective and literal types. The main questions motivating the experiments are:

-How do different literal and planar techniques perform in simple multi-display environments?

-How do planar and perspective techniques compare in complex multi-display environments?

5.2. Experiment 3: Literal vs. Planar

Experiment 3 was chronologically the first of all the studies contained in this dissertation. It was the first published multi-display study to quantitatively compare a broad set of interaction techniques for reaching small and large distances across displays. At the time of the design, perspective techniques had not been identified as a separate group, and therefore are not included here. However, the selected techniques do include several types of planar and literal CDOM techniques, and results are therefore pertinent to this chapter. The experiment tests two main scenarios: within hand's reach destinations (i.e., locations that can be directly reached without having to walk), and beyond hand's reach destinations (i.e., locations that fall outside the immediate reach of the hands of the user). Each one of the scenarios is tested in a different part of the experiment, and the details of each part of the experiment are reported separately.

5.2.1. Goals

The main goal of this experiment was to compare a relevant set of interaction techniques (described below in Section 5.3.3) in terms of:

- Efficiency (how fast a task can be completed with each technique)
- Accuracy (how often errors were made with each technique)
- User preference (what techniques were preferred by the participants of the experiment)

5.2.2. Apparatus

A multi-display system was set up with two computers: a pen-activated tablet PC with a 15.5cm by 21cm screen (768x1024 pixels) and a desktop PC controlling two projectors capable of projecting an image of 121 by 81cm (1536x1024 pixels) combined. The projectors were mounted vertically to project onto a large white table (151x178cm) that incorporated a pen-driven digitizer also connected to the desktop PC. Both the digitizer and the tablet PC were pressure sensitive. Figure 35 depicts the experimental setup. In our experiment, the tablet PC was fixed into position so that the software could always provide a coherent visual output. In the *within hand's reach* part of the experiment, the tablet PC was attached to the table on the side closest to the projected surface (Figure 35.A) and in the *beyond hand's reach* part of the experiment to the opposite side (Figure 35.B). In the second part the user had visual feedback only in the projected surface area at the far end of the table. The two PCs ran separate programs that communicated with each other through TCP sockets over an ad-hoc wireless network.

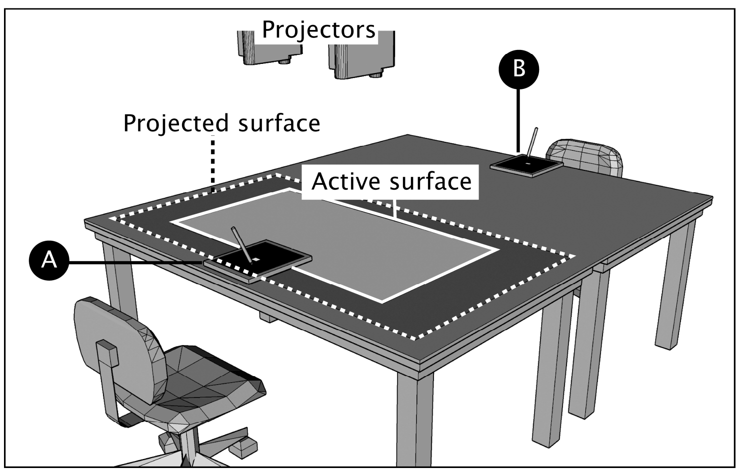

Figure 35. The experimental setup. In the within hand's reach experiment the tablet PC was placed at position A and in the next experiment it was placed at position B.

5.2.3. Techniques

Due to practical limitations in the number of techniques that can be compared, we picked a group of prominent techniques from the research literature that were different from each other. The six techniques that we compared are: Pantograph, Slingshot, Pick-and-Drop, Corresponding-Gestures, Radar View, and Press-and-Flick.

Pantograph, Slingshot and Press and Flick fall into the planar category; Pick-and-Drop and Corresponding-Gestures are literal techniques. Radar View can be considered literal or planar, depending on whether the miniature is sufficient for the manipulation of the objects (see discussion at the end of Section 4.1.3). However, for the targeting task that was tested, the Radar View can be safely considered a literal technique.

Accompanying video figures illustrate instances of interaction with each one of the techniques (the files are named after the techniques).

5.2.3.1. Pick-and-Drop

The Pick-and-Drop technique was implemented as defined by Rekimoto [1997]. The user can select an object by touching it with a digital pen and then lifting the pen. The selected object then disappears and is visible only when the pen is moved close to the surface. As the user touches the surface again the object is dropped at that location. Pick-and-Drop is a literal technique since the input model of the technique and the physical configuration overlap perfectly.

5.2.3.2. Corresponding-Gestures

Corresponding-Gestures is similar to Pick-and-Drop, but instead of touching the surface to select and deselect an object, the user makes a special predefined gesture. To select the object the user draws the gesture starting inside the object (see Figure 36) and then lifts the pen. When the object is selected, it is visible only when the pen is moved close to the surface. To deselect (drop) the object the user draws the same gesture, and the object is then dropped so that its centre coincides with the starting point of the gesture (see Figure 36). This technique was inspired by [Rekimoto et al., 2003; Holmquist et al., 2001; Hinckley et al., 2004], but it is a literal technique for the same reasons as Pick-and-Drop.

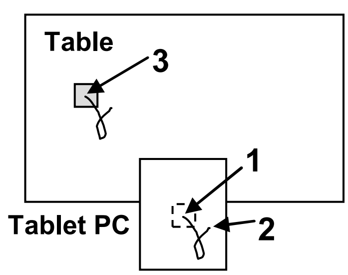

Figure 36. Corresponding-Gestures. 1) Starting point of the selecting gesture, 2) end point of the selecting gesture, 3) dropping gesture.

5.2.3.3. Slingshot

This technique is based on the metaphor of a physical slingshot and is similar to the Drag-and-Throw technique [Hascoët et al., 2003]. It was implemented as follows: the pen touches the object, then without losing contact with the surface the pen is moved backward and released. The object is 'thrown' forward proportionally to the size of the pen's backward stroke. The dependence is linear. In order to keep the effective distance moved by the pen the same for the Pantograph, Slingshot and Radar, the amplification factors were equal across these techniques for each part of the experiment. In the within hand's reach experiment, the amplification coefficient was equal to 15, in the beyond hand's reach it was 40.

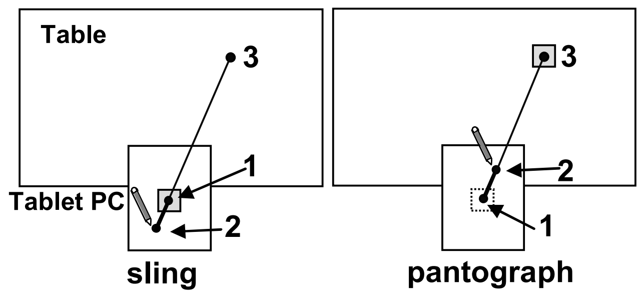

Figure 37. Slingshot and Pantograph. 1) initial position of the pen, 2) current position of the pen, 3) destination of the object.

While moving the pen backwards the user also can move it to the right and left to define the direction of the shot. The continuous visual feedback shows a line indicating where the object will end up if the pen is released (see Figure 37). The figure corresponds to the within hand's reach experiment; in the beyond hand's reach experiment, there was a feedback-blind area between the tablet and the projected surface (see Figure 35). This holds also for the Pantograph and Press-and-Flick techniques. This technique belongs to the planar group, since the display surfaces in the physical space are considered part of a single plane by the interaction technique.

5.2.3.4. Pantograph

The implementation of the Pantograph technique is similar to the Push-and-Throw technique [Hascoët et al., 2003]. Here, the short movement of the pen is mapped into the long movement of the object. Therefore this technique is very similar to the Slingshot, but instead of a backward movement it uses forward movement. In the within hand's reach experiment the amplification coefficient was equal to 15, in the beyond hand's reach, 40.

The user selects an object by touching it, then without lifting the pen drags in the desired direction. As with the Slingshot, a feedback line shows the destination. The further the object is moved from the initial position, the further it will be thrown in the same direction (Figure 37).

Pantograph is a planar technique for the same reasons as Slingshot.

5.2.3.5. Press-and-Flick

In this technique the throwing distance for an object is defined by pen pressure, and the direction is defined by a single straight stroke toward the desired target. These actions are separate. The user first defines the distance (by pressing the pen against the surface) and then moves the pen slightly from the starting point to fix the throwing distance. This is signaled to the user by changing the color of the circle. At this point, only the direction will now be measured (see Figure 38). The dependence between the pressure and the throwing distance was a monotonic function adjusted using pilot testing and chosen based on the guidelines suggested in [Norman, 2002]. This technique also uses a planar model for the combination of displays, although the first part of the gesture (the pressure) uses pressure instead of distance to modify the reach range.

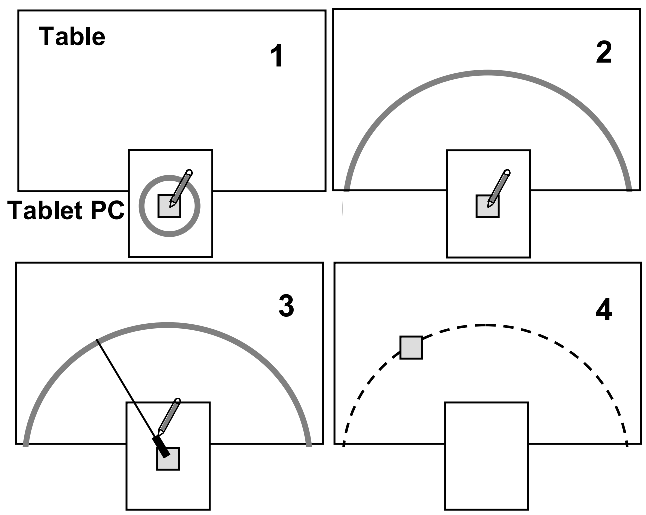

Figure 38. Press-and-Flick. 1) Pen touches the surface (low pressure); 2) user increases the pressure; 3) user starts to move the pen - the distance is automatically fixed, the color of the circle changes and the line which indicates where the object will end up is displayed; 4) pen is released.

5.2.3.6. Radar

The Radar (or Radar View) technique uses a miniature representation (a map) of the surrounding environment (i.e., the tabletop display and the tablet PC). When the pen touches the object the map appears. The map is placed so that the position of the pen is the same in both representations (see Figure 39). The map provided is similar to that in Drag-and-Pick [Baudisch et al., 2003] but allows continuous positioning without distorting the shared space. When the user starts to move the pen (without lifting it from the tablet) a small line connects the starting point of the stroke with the actual position of the pen's tip. The place on the map where the pen is released determines the final position of the object (see Figure 39).

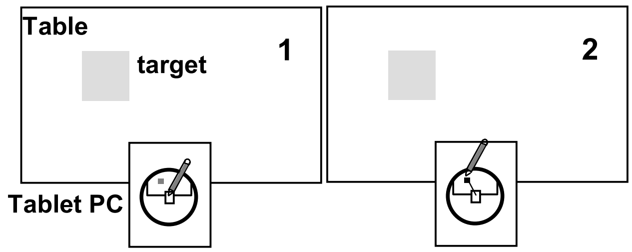

Figure 39. Radar. 1) Pen touches the object, a reduced representation (map) of the surrounding environment appears, 2) user moves the pen to the representation of the target within the map and lifts the pen.

The Radar View is a planar representation of the physical space; however, if the radar itself is detailed enough to perform the whole manipulation, it becomes a single-display, direct input technique (see discussion at the end of 4.1.3). Since this is the case for the task of the experiment, I will consider the Radar View here as a special case of a literal technique.

5.2.4. Participants

The two parts of the experiment (within hand's reach and beyond hand's reach) were conducted with different sets of participants. A total of 18 participants with ages ranging from 18 to 44 took part in either one of the two parts of the experiment (never both). Participants were recruited amount the population of the University of Saskatchewan and compensated with CAN$10 for their time. All participants read and signed a consent form and a demographic data form at the beginning of the session, and a post-study questionnaire after all tasks were completed. Samples of the consent form and the questionnaire are reproduced in Appendix A. The experiment and the handling of participants were carried out according to the University of Saskatchewan's ethical guidelines for experiments with human subjects.

5.2.4.1. Within Hand's Reach

The first part of the experiment was conducted with 10 right-handed participants (3 females, 7 males). All participants had previous experience with graphical interfaces. The experiment took about one hour to complete, including the filling of the forms, basic explanations, training, rest in between the conditions, and the experimental tasks themselves.

5.2.4.2. Beyond Hand's Reach

This part of the study was conducted with 8 right-handed participants (3 females and 5 males). All participants had previous experience with graphical interfaces. The experiment took about an hour to complete, including filling of the forms, basic explanations, training, rests, and the experimental tasks.

5.2.5. Task

Each subject was instructed to move several objects from the tablet PC to various targets on the table, using the different interaction techniques. The object to be moved was a 15x15mm graphical icon, and the targets were 80x80mm squares placed in the projected space of the table.

The trial was considered a 'hit' if the object ended up completely inside the area of the target, and a 'miss' otherwise. The system provided distinctive sound feedback for hits and misses. The subjects were asked to perform repetitions of this task using the different interaction techniques, as fast as possible but without missing the target.

The experimental software measured completion time from the moment that the object appeared on the screen until the object was released in the target. It also measured the number of errors (the participant dropped the object outside of the target).

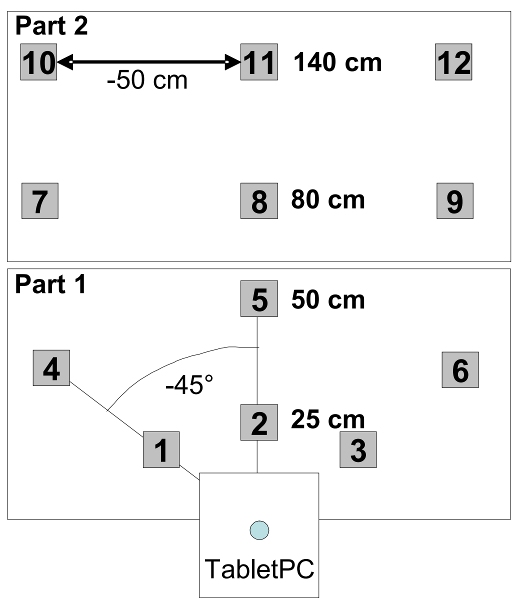

Figure 40. Target locations in both parts of the experiment.

The within hand’s reach part of the experiment tested tasks involving targets one to six as displayed in Figure 40, which appeared in an unpredictable order. The tasks of the beyond reach part of the experiment required reaching targets seven to twelve.

5.2.6. Questions

The study was structured to find out:

- How do techniques compare to each other in terms of speed
- How do techniques compare to each other in terms of accuracy
- Which techniques are preferred by participants

5.2.7. Experimental Design

Both parts of the experiment shared a common experimental paradigm and differed only in the location of the targets and the technique tested. It did not make much sense to test literal techniques in the beyond hand’s reach part of the experiment because that would force participants to stand up and walk towards the target. We do not need to set up an experiment to find that completion times for these techniques in the beyond hand’s reach of the study would be long compared to the other techniques. Therefore literal techniques (except for the Radar View) were not tested in the second part of the study. This made time for some extra repetitions of the task, and increased statistical power, which was then used to reduce the number of participants to eight.

All participants first received a demonstration of all techniques, after which they performed all the tasks with all techniques. For each technique, participants carried out training first (20% of the trials), and then took a short break and started with the real trials. The order of the different techniques was randomized across participants using a random Latin square to prevent bias due to order effects.

5.2.7.1. Within Hand’s Reach

The experiment used a 6x2x3 within-participants factorial design with a variety of planned comparisons. The factors were:

- Technique (Radar, Pantograph, Pick-and-Drop, Corresponding-Gestures, Press-and-Flick, Slingshot)
- Target distance (25cm, 50cm)
- Target angle (-45˚, 0˚, +45˚)"

For each technique and location, participants completed two training trials and eight test trials, for a total of 72 training trials and 288 test trials.

At the end of the experiment participants were asked in an exit questionnaire to rank the six techniques in order of preference.

5.2.7.2. Beyond Hand's Reach

The experiment used a 4x2x3 within-participants factorial design with planned comparisons. Factors were:

- Technique (Radar, Pantograph, Press-and-Flick, Slingshot)
- Target distance (80cm and 140cm)
- Target side (-50cm, 0 cm, 50cm, where 0cm is directly forward and positive numbers are to the right of the participant)

The target side locations were chosen such that they are close to the target angles of the previous experiment and still within the projected space. For each location and technique, participants completed two training trials and ten test trials (for a total of 240 test trials and 48 training trials).

At the end of the experiment participants were asked in an exit questionnaire to rank the four techniques in order of preference.

5.2.8. Results

5.2.8.1. Within Hand's Reach

Some trials were deleted when participants accidentally released the pen from the object, causing a series of errors. In techniques like Pick-and-Drop, and Corresponding-Gestures, the users sometimes repeated their gesture to re-pick the object causing it to be dropped close to the initial point. In the case of Press-and-Flick, in an attempt to change the pressure the user sometimes removed the pen from the surface causing the object to be dropped close to the initial point. These errors were due to the implementation rather than the design of the techniques, and so were removed from the data. Fewer than 15 trials per subject were deleted. The distribution of deleted trials over various interaction techniques is shown in Table 3 (numbers between parentheses).

	Radar	Pantograph	Slingshot	Press-and-Flick	Pick-and-Drop	Corresponding-Gestures
1	0.0%	1.3%	1.3% (1)	4.1% (3)	0.0% (4)	5.5% (3)
2	0.0%	2.6%	2.6%	5.4% (2)	1.3% (3)	6.9% (3)
3	0.0%	2.6%	6.7%	2.6% (1)	1.3% (4)	1.3% (2)
4	1.3% (1)	1.3%	3.9%	16.7% (3)	0.0% (5)	4.1% (4)
5	0.0%	5.3%	3.9%	11.1%	1.3% (1)	2.7% (5)
6	1.3%	3.9% (1)	3.9%	4.0% (2)	0.0% (2)	2.7% (3)

Table 3. Error rates in the within reach part of the experiment corresponding to target locations numbered in Figure 40. The numbers in parentheses indicate the trials that were discarded due to inadvertent errors in releasing the pen.

An ANOVA test was performed with interaction technique as main factor, controlling for the location of the target (distance: 25 or 50cm, and angle: -45° or 0° or 45°), and modeling the participant as a random factor. The results indicate a clear effect of interaction technique ($F_{5,45}$ = 83.8, p <.001) and, as expected, of the location of the target (angle: $F_{2,18}$ = 4.2, p <.033, distance: $F_{1,9}$ = 56.2, p <.001). The interactions *distance-interaction technique* and *angle-interaction technique* were also significant ($F_{5,45}$ = 17.6, p <.001 and $F_{10,90}$ =4.3, p <.009 respectively), but the effects are in the same direction as can be seen in Figure 41, Figure 42 and Table 4.

Post-hoc pair-wise comparisons (Tamhane T2, unequal variances) of interaction techniques yielded significant differences (all p <.001) in trial completion times for all pairs except for the pair *Pantograph / Synchronized Gestures* in the far targets. The techniques can be ranked in decreasing order of performance (for the two distances) as shown in Figure 43.

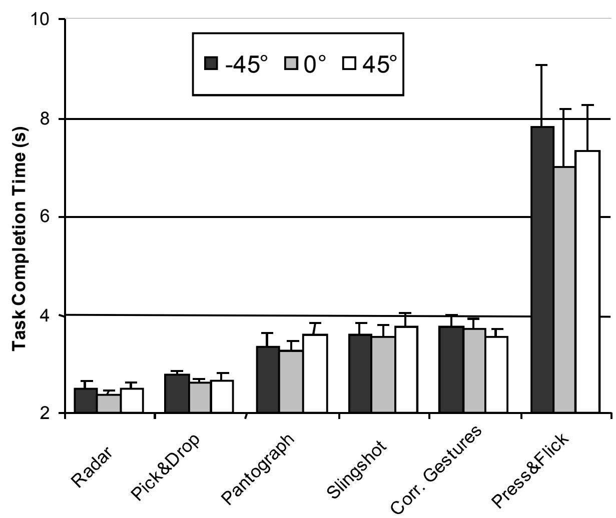

Figure 41. Mean trial completion times for different interaction techniques in the within reach part of the experiment (25 cm). Note that the vertical axis starts at the 2s mark.

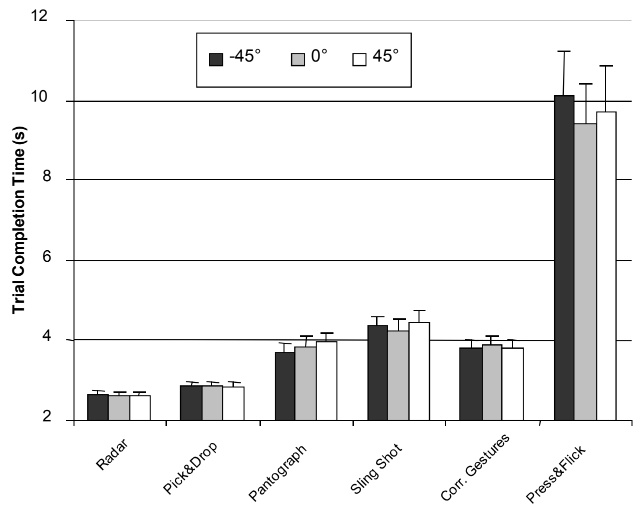

Figure 42. Mean trial completion times for different interaction techniques in the within reach part of the experiment (50 cm). Note that the vertical axis starts at the 2s mark.

Table 4. Average completion times for the different techniques (rows) in the different angles (columns) for the two distances of the within reach part of the experiment (25cm and 50cm). Numbers in parenthesis represent standard error.

	25cm			50cm		
	Left	Center	Right	Left	Center	Right
Radar	2.52 (.05)	2.40 (.03)	2.48 (.06)	2.67 (.03)	2.61 (.03)	2.61 (.04)
Pick&Drop	2.77 (.04)	2.62 (.04)	2.67 (.05)	2.89 (.03)	2.89 (.03)	2.85 (.04)
Pantograph	3.38 (.10)	3.29 (.08)	3.61 (.08)	3.74 (.08)	3.90 (.09)	4.01 (.07)
Slingshot	3.63 (.09)	3.54 (.09)	3.86 (.10)	4.38 (.09)	4.26 (.11)	4.50 (.11)
Corr.Gestures	3.80 (.09)	3.69 (.08)	3.56 (.07)	3.82 (.07)	3.88 (.08)	3.81 (.07)
Press&Flick	7.78 (.47)	7.06 (.44)	7.32 (.34)	10.25 (.45)	9.42 (.39)	9.89 (.41)

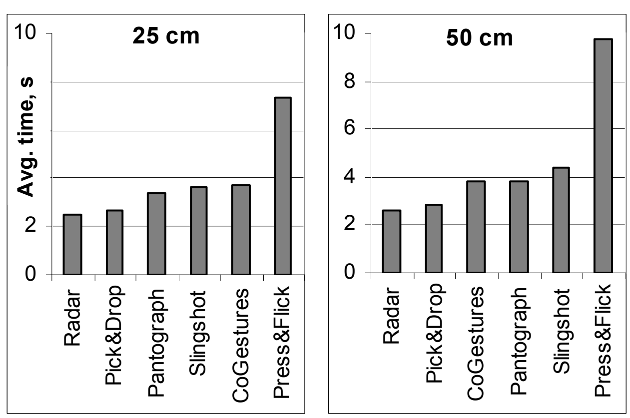

Figure 43. Speed ranking of the interaction techniques in the within reach part of the experiment.

The number of misses was significantly different between the different interaction techniques in both the 25 cm range ($\chi^2_{5,\ 13}$=16.784, p<0.001) and the 50cm range ($\chi^2_{5,\ 13}$)=35.895, p<0.001). In both the 25 cm and 50 cm ranges, Radar and Pick-and-Drop had significantly fewer errors than either Press-and-Flick or Corresponding-Gestures (see Table 3).

We further asked people to rank the techniques in terms of which technique they felt was fastest and which they preferred overall. Results for these two questions were very similar: most people ranked the Radar highest (7 of 10 people), with a few preferring either Pick-and-Drop (2 people) or Corresponding-Gestures (1 person); in addition, all participants ranked Press-and-Flick last. Ordered by average rank, the techniques are Radar, Pick-and-Drop, Corresponding-Gestures, Pantograph, Slingshot, and Press-and-Flick. These rankings matched the efficiency ranking of Figure 43.

5.2.8.2. Beyond Hand's Reach

Again, fewer than 15 trials per participant were discarded due to inadvertent errors in releasing the pen. The distribution of deleted trials over various interaction techniques is shown in Table 5 (numbers between parentheses).

	Radar	Pantograph	Slingshot	Press-and-Flick
7	6.7%	1.3% (1)	1.4% (7)	1.4% (8)
8	1.3%	9.7% (1)	9.6%	15.6% (6)
9	0.0% (2)	5.6% (4)	17.9% (1)	1.4% (7)
10	6.7%	12.9% (1)	15.9%	37% (10)
11	8.1%	9.6%	17.6%	20.6% (4)
12	*2.6%*	*9.6%*	*9.6%*	*27.6% (6)*

Table 5. Error rate for all interaction techniques in the beyond reach part of the experiment corresponding to target locations numbers in Figure 40. The numbers in parenthesis indicate the trials that were discarded due to inadvertent errors in releasing the pen.

An ANOVA test was performed with interaction technique as main factor, controlling for the location of the target (distance: 80cm or 140cm, side: -50cm, o or 50cm), and modeling the participant as random factor. The results indicate also a clear effect of interaction technique ($F_{3,21}$ = 122.9, p <.001) and also of the side of the target location ($F_{2,14}$ = 43.6, p <.001), but not of the distance ($F_{1,7}$ = 0.092, p = .77). This time only the statistical interaction between the interaction technique and the side was significant ($F_{6,42} = 2.24$, $p < .024$). However, as in the within reach part of the experiment, the effects were in the same direction (see Figure 44).

Post-hoc pair-wise tests (Tamhane T2, unequal variances) were all significant except for the difference between Pantograph and Slingshot ($p = .32$); the pair-wise tests yield the ranking reflected in Figure 45.

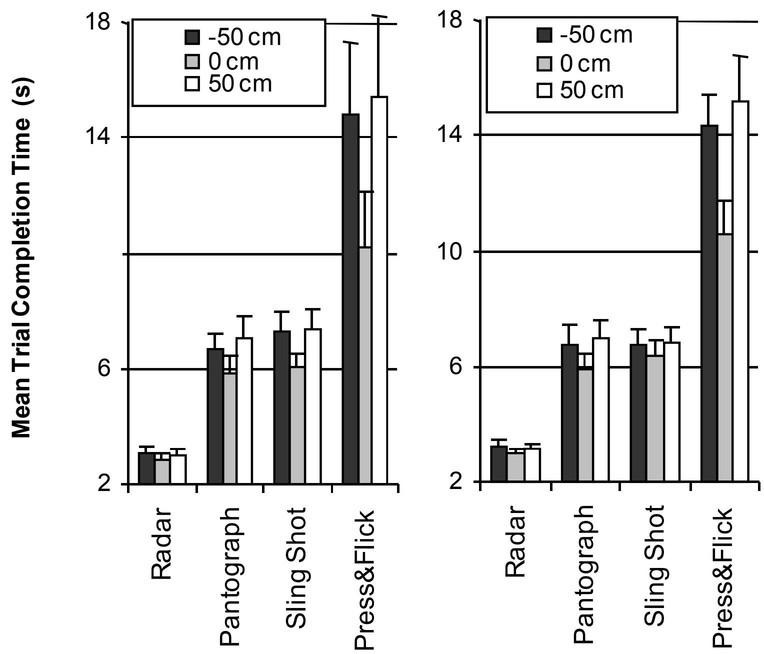

Figure 44. Mean trial completion times for different interaction techniques in the beyond reach part of the experiment, 80 (left) and 140 (right) cm. Note that the vertical axis starts at the 2s mark.

Table 6. Average completion times for the different techniques (rows) in the different angles (columns) for the two distances of the beyond reach part of the experiment (80cm and 140cm). Numbers in parenthesis represent standard error.

	80cm			140cm		
	-50cm	0cm	+50cm	-50cm	0cm	+50cm
Radar	3.12 (.07)	2.86 (.06)	3.03 (.07)	3.23 (.07)	3.01 (.06)	3.14 (.06)
Pantograph	6.70 (.19)	5.88 (.24)	7.02 (.29)	6.76 (.25)	5.91 (.20)	6.87 (.23)
Slingshot	7.24 (.24)	6.09 (.18)	7.32 (.27)	6.81 (.21)	6.38 (.21)	6.77 (.20)
Press&Flick	14.91 (.96)	10.37 (.79)	15.52 (1.06)	14.46 (.60)	10.35 (.41)	15.13 (.62)

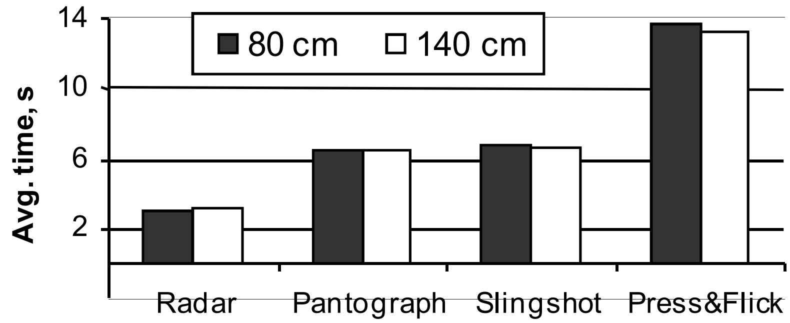

Figure 45. Mean trial completion times for the beyond reach part of the experiment. Averaged across direction.

The overall success rate was lower for this part of the experiment: 1690 hits and 164 misses overall (91.2%). The number of misses were significantly different between different interaction techniques in both the 80 cm ($\chi_{23,9}$ = 23.579, p <.001) and the 140cm ranges ($\chi_{23,9}$ = 30.882, p <0.001). Pair-wise comparisons showed that Radar had significantly more hits than either Press-and-Flick or Slingshot for the 140cm range, and that Pantograph had significantly more hits than Press-and-Flick.

We also asked for people's preferences in the second part of the experiment, and again perception of speed and preference were very similar. The Radar was again ranked first most frequently (7 of 8 people), and Press-and-Flick was ranked last by all participants. In order of mean rank, the techniques are: Radar, Pantograph, Slingshot, and Press-and-Flick.

5.2.9. Discussion

Figure 46 summarizes the effect of target distance on trial completion times for the different interaction techniques of both experiments. Figure 40, Table 3 and Table 5 show the target positions used in both the experiments and the corresponding miss and hit rates for each interaction technique. Techniques that were faster were also more accurate. Radar was the most accurate interaction technique while Press-and-Flick was the least accurate. Below we discuss the performance of the different techniques and how this relates to the group they belong to.

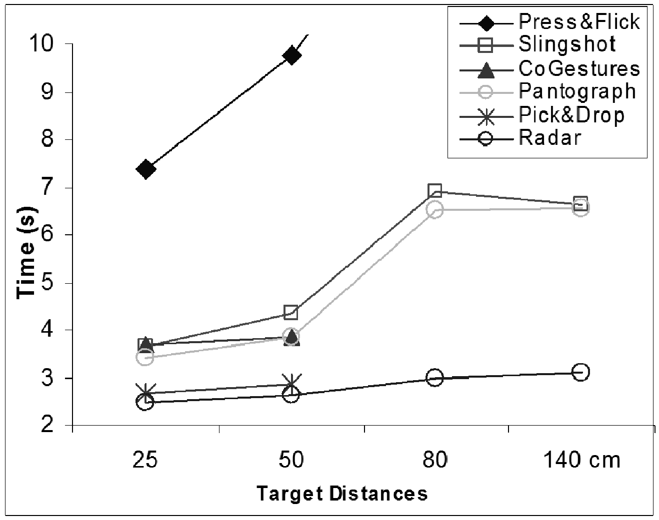

Figure 46. Summary of completion times for different interaction techniques according to target distance. Note that the vertical axis starts at the 2s mark.

5.2.9.1. Literal vs. Planar Techniques

The two techniques that showed the best performance in the study (Radar and Pick-and-Drop) fall in the category of literal techniques. In the next performance group we find the planar techniques (Pantograph and Slingshot), which are equivalent in terms of input to the Radar, but showed inferior performance. The difference in performance between Radar and Pantograph is explained by the location of the feedback is displayed. Since Radar in this study is a literal technique, feedback and input overlap, whereas in the planar technique there is a disconnect (indirection) between the two. This result indicates that the coupling between the input and the output typical of literal techniques represents an advantage in performance.

The two literal techniques (Radar and Pick-and-Drop) are slightly different from each other, but so is the amplitude of the gesture required. This performance difference does not completely conform with Fitts's Law [Fitts, 1992; MacKenzie, 1991] because the index of difficulty remains the same (ID does not change if amplitude of the gesture and size of the target are scaled by the same factor). However, the difference is very small, and may be caused by the difference in muscular groups used for a small gesture (which would typically only involve hand and wrist movement), and a large gesture (that involves the whole arm).

It is important to note that even though the Pick-and-Drop technique performed slightly worse than Radar, it is still much faster than the best planar techniques. Accuracy results follow the same pattern as completion time results. These results indicate that the differences in performance due to the type of technique (the focus of this chapter) are more important than subtle changes in the implementation of the technique (such as the amplitude of the movement required).

5.2.9.2. Radar View

Radar View can be considered a literal technique in this study because the task did not require any real interaction with real targets; instead the task could be completed by looking only at the miniature (targets and displays would be visible in the miniature). In most situations this is not realistic, since the destination of an object is not necessarily known by the system in advance, and users would have to find the destination in the physical space, then figure out the corresponding point in the miniature, and finally execute the movement.

It is also not clear whether radars will perform well if they are not oriented and positioned with respect to the user [Kortuem et al., 2005]. If orientation and location are required, Radar or world-in-miniature techniques would suffer from most of the problems of perspective techniques, but with the drawbacks discussed in the previous paragraph.

In general, existing world-in-miniature techniques (including Radar) could be considered planar because the representation of space is flat; however, it would be possible to conceive a world-in-miniature technique where the miniature would represent the environment as it looks from the point of the user. This technique could be considered a perspective technique.

The unanswered questions about radar techniques indicate the need for further study on the topic. However, the issues involved in world-in-miniature techniques far exceed the scope of this dissertation, and have therefore been left as future work.

5.2.9.3. Pantograph vs. Slingshot

Pantograph and Slingshot are only different in how the spatial gestures are mapped. Although the differences in performance between the two were very small, the advantage of the Pantograph indicates that forward mappings (the object moves forward if the gesture goes forward) are preferable to inverted mappings. This result is consistent with Stimulus-Response Compatibility studies discussed in Sections 3.1.4.1 and 4.3.

5.2.9.4. Pick-and-Drop vs. Correspondin- Gestures

Corresponding-Gestures had the poorest performance within the literal techniques. This result was expected because, other than the extra gesture required for the object selection and object release, this technique is equivalent to Pick-and-Drop. The extra gestures can only hinder performance. However, the extra gestures also result in advantages for the implementation and in collaborative scenarios. For example, it might be expensive to get all the displays to work with the same touch technology and to make them recognize different input devices. This problem is solved with the Synchronized Gestures technique because the information about the identity of the device or the user can be encoded in the shape of the gesture, without requiring specialized hardware. This characteristic also allows different people to perform the same gesture, effectively making the transfer of objects between people possible (as long as they have a means to agree on the type of gesture that they will perform). In other words, Synchronized Gestures has added power and implementation flexibility, but at the cost of performance.

5.2.9.5. Press-and-Flick

Press-and-Flick was consistently and by far the least efficient technique. The two-phase interaction, the unintuitive relationship between pressure and distance and the poor control of pressure with the pen used in the experiment are possible reasons for the inferior performance. Pressure based input devices need to be studied carefully to make them easier to control. These observations complement the findings of Ramos and colleagues [Ramos et al., 2004].

The poor performance of this technique might also reflect the poor mapping between dimensions. Pressure and distance do not seem to have much dimensional overlap, and this probably affected the results.

5.2.9.6. Limitations of the Experiment

In our experiments the users were required to perform the task as fast as possible and with strict accuracy requirements (i.e., pre-determined target sizes and specific well-identified locations). In a real world MDE, performance requirements are often not so strict. It is not always the aim of the user to send an object as fast or as accurately as possible. However, there are no obvious reasons to believe that changes in the scenario will change the performance relationships between the techniques.

With regard to the applicability of the results this experiment only presents a very simple MDE, where the only discrepancy between the two displays involved is their size. Therefore, extrapolating these results to situations with more displays has to be done with care. It is likely, however, that an irregular MDE with more displays would only increase the performance gap between literal and planar techniques (in the scenarios in which these are comparable, i.e., within hand's reach).

5.3. Experiment 4: Perspective vs. Planar

Experiment 3 compared two of the three categories relevant to this chapter: literal and planar; although it tested 6 techniques, none of them can be considered perspective techniques. Perspective techniques were not tested because at the time of the design of experiment 3, perspective techniques had not been identified as a different group, and these were only starting to be published (e.g., [Ashdown et al., 2005; Benko and Feiner, 2005, 2007]).

Experiment 4 tests perspective techniques against the most common planar technique: Stitching. As discussed in Section 4.1.2, perspective techniques are interesting because they use knowledge about how the space is perceived by the user, and might surpass the performance of planar techniques without losing the power of acting remotely (as with literal techniques).

This experiment was designed to test whether perspective techniques actually provide any advantage over non-perspective techniques. The experiment does not include any literal technique in the comparison because, by their very nature, these do not afford access to remote space, something that is potentially very limiting in most multi-display scenarios. In certain cases it might be interesting to know whether walking up and touching an object (literal) is comparable to acting remotely on it, and at what threshold distance it is no longer worth it to move in order to physically reach the object. However, this question is less relevant than the one answered by this experiment because we already have some evidence that indicates that users generally prefer not to have to walk or change their posture to act on an object [Voida et al., 2005], and that, even when already standing next to a display, a distance of more than 4.5 feet (1.3 meters) results in a large performance penalty that surpasses the benefit of direct input [Khan et al., 2004].

Experiment 4 tests three techniques: Stitching, Laser Pointing, and Perspective Cursor. Stitching is the most common planar technique for multi-display environments, and probably the most common CDOM technique overall. The Laser Pointer technique is likely the most common perspective technique in large displays and multi-display environments. Perspective Cursor is our own invention, and makes perspective and planar techniques more comparable because the input device is common with Stitching.

Testing two perspective techniques will also allow us to measure how much of the performance difference can be explained through changes in the input device.

5.3.1. Goals

The main goal of this experiment is to compare the two perspective techniques and the planar technique and establish:

- Which technique performs better than the others and in which conditions
- Which techniques are preferred by users

5.3.2. Apparatus

For the experiment, we developed a prototype single-user multi-display environment through which we can test most kinds of cross-display transitions of the cursor. The setting consists of three fixed displays and one mobile display. The three fixed displays are a large vertical wall-projected screen, a projected tabletop display and a regular flat monitor. The mobile display is a tablet PC.

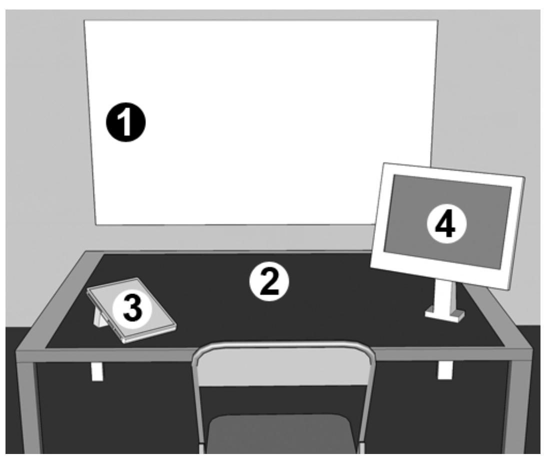

Figure 47. Experimental setting. 1) wall display 2) table-top display 3) tablet PC 4) flat screen.

Figure 47 shows the physical locations of all the elements. The tabletop display (2) is a projected table with an image of 1024x768 pixels and 158x124.5cm in size. The wall display (1) has the same resolution but the projection is slightly smaller (136x101.5cm). The flat screen (4) is a 15” LCD monitor with a resolution of 1024x768 pixels. The tablet PC’s display (3) is 15.5x21cm with an image of 768x1024 pixels.

We use a total of three computers to control all the displays. The main application resides in a Pentium IV PC that also controls the two big displays. The flat panel and the tablet PC are controlled by independent machines connected to the main application by a dedicated Ethernet network.

For relative-input we use a wireless mouse. Position tracking is provided by a Polhemus® Liberty tracker with three tethered 6-DOF sensors. One sensor is attached to a baseball cap that measures the user's head position, another is attached to the tablet PC, and one, in the shape of a pen with a button, serves as the virtual laser pointer.

The system kept an updated 3D model of the whole setting, including the displays, the position of the user's head, the position and orientation of the pen (laser pointer) and the position and orientation of the mobile display. Although the tracking technology used is affected by metallic and magnetic objects, the setting was designed so that accuracy of tracking was not an issue, except for the case of the tablet PC when using Laser Pointer, which will be discussed later.

5.3.3. Techniques

Three techniques were implemented in this prototype for the experiment. Of the three, one is planar (Stitching) and two represent perspective techniques (Laser Pointer and Perspective Cursor). The accompanying video figures show interaction instances with all three techniques (files are named accordingly).

The following sub-sections describe the implementation of the techniques.

5.3.3.1. Stitching

In this technique, the movement of the mouse is related to the changes in the coordinates of the cursor in a linear fashion. When the cursor reaches the border of a particular display, the system checks if there is another display assigned to that end; if so, the cursor continues moving onto the new display. If there is no other display stitched to that end, the cursor just stays at the border (hard edges).

Figure 48 depicts the actual stitching implemented for our setting, which was designed to be as close as possible to a "flattening" of the room's 3D model into a 2D map.

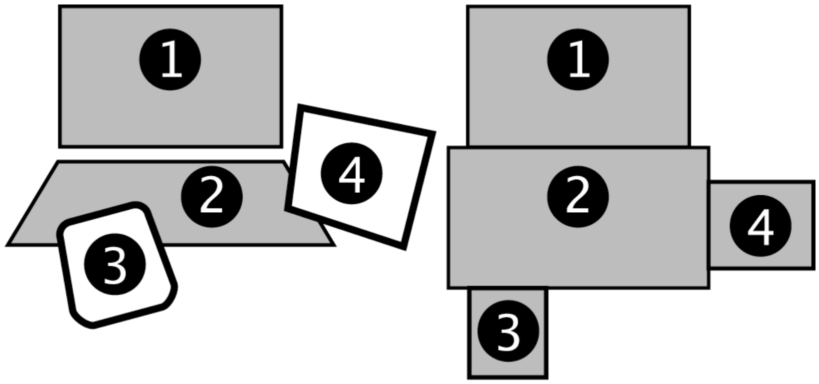

Figure 48. Multi-display environment and its 2D stitching of control spaces.

The C/D ratio in centimeters is kept constant, meaning that the small displays take less mouse movement to cross, although the number of pixels is the same.

5.3.3.2. Laser Pointing

Our implementation of Laser Pointing uses a 6-DOF sensor in the shape of a pen with a button close to the tip. To obtain the position of the cursor we virtually intersect a mathematical line coming from the tip of the pen in the longitudinal direction with the virtual model of the room. If there is a display in the way, the two-dimensional coordinate point of the intersection relative to that display is considered the current position of the pointer. If the line intersects more than one display, the display closest to the pen is chosen as the one displaying the cursor. When the line does not intersect any display, nothing is shown.

In short, the pen works as a laser pointer but for the fact that it controls the system's pointer instead of a red dot, and that it does not display anything when pointed to a space without displays in the way. The button in the pen generates the same kind of events as does a mouse button.

Due to technology constraints, the pen could not be too close to the tablet PC without appreciable distortion (distortion appeared at distances of around 4cm). All subjects of the study were instructed not to bring the pen too close to the tablet PC to avoid this effect.

Laser Pointing is a perspective technique because the degree of control varies depending on the position of the device, which is dependent on the position of the person (see Section 4.1.2). In terms of general pointing performance, Laser Pointing has been found to be inaccurate and slow [Myers et al., 2001; Oh and Stuerzlinger, 2002]. Laser Pointing and Stitching can be compared in terms of performance but, if this is the only comparison, we will not be able to tell whether the differences come from the different nature of the technique (planar vs. perspective) or the input device (pen vs. mouse), leaving one of the main questions of this chapter unanswered.

5.3.3.3. Perspective Cursor

Perspective Cursor is a novel technique that I devised to have the advantages of both remote pointers and the mouse without some of their most obvious drawbacks. This section describes in detail how Perspective Cursor works.

Perspective Cursor uses a relative positioning input device (e.g., a mouse or a trackball) together with the user's point of view to determine how displays are located in the field of view. Since the main input device is the mouse, we compare Perspective Cursor to the planar techniques (Stitching) without the possible confound of the input device because both use the same device.

Perspective Cursor works as follows. We obtain in real time the 3D position coordinates of the head of the user (but not the orientation or the gaze direction) and at the same time, we maintain a three-dimensional model of the whole environment (including the actual position of all the screens). The model, together with the point-of-view coordinate of the user's head, lets us determine which displays are contiguous in the field of view, something very different to displays actually being contiguous in 3D space (Figure 49).

The position and movement of the pointer is calculated from the point of view of the user, so that the user perceives the movement of the pointer across displays as continuous, even when the actual movement of the pointer considered in three dimensional space is not. Figure 49 shows several inter-display transitions that illustrate how the pointer moves between two displays when the displays are A) contiguous from the point-of-view of the user, B) overlapping from the point-of-view of the user and C) separate from the point-of-view of the user.

As can be observed in Figure 49.C, the pointer travels through empty space to get from one display into the next. Actually, the cursor can be in any position around the user, even if there is no screen there to show the graphical representation of the cursor.

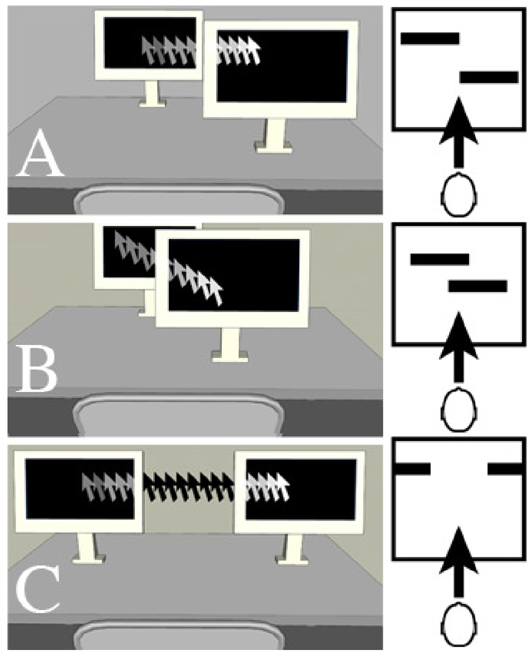

Figure 49. Examples of display transitions of Perspective Cursor. A) The displays are in different planes, but appear contiguous to the user. B) Displays that overlap each other. C) The cursor travels across the non-displayable space to reach the other display (the black cursors are only illustrative)

Perspective Cursor uses the users' head position (but not orientation) as the origin of the intersecting line discussed above. The orientation of the line is determined by the movements of the mouse so that a vertical movement of the mouse results in an increase or decrease of the angle of the line with respect to the equator (or the horizontal plane). Conversely, a horizontal movement of the mouse changes the longitudinal orientation of the line (i.e., the angle with respect to the vertical plane). Wherever the virtual line intersects the surface of a display, there lies the Perspective Cursor.

Alternatively, Perspective Cursor can be explained as a virtual laser pointer located at the eyes of the user and which orientation is controlled by the movement of the mouse (changes of the x mouse coordinates change yaw, and y changes change pitch).

The cursor keeps a constant size relative to the user, i.e., the image of the cursor varies in size and shape depending on the position and orientation of the surface where it is being displayed, but it projects the same image on the user's retina. For the experiment, the size of the cursor was calculated to be about three times the size of a normal cursor seen in a 1024 by 768 screen at a normal viewing distance (40cm), covering an angle of around 2 degrees (angular size of the cursor was multiplied in the video figure for illustration purposes).

If the position of the head changes but the mouse is not moved, the cursor stays in the same place of the same display. To prevent users from losing Perspective Cursor in inter-display (blank) space, we used a variant of the Halo technique [Baudisch and Rozenholtz, 2003] adapted for 3D environments.

There are not many environments in which the users are completely surrounded by displays, meaning that users might lose the pointer in non-displayable space. The solution that we implemented is a perspective variant of Halo [Baudisch and Rozenholtz, 2003]. Halos are circles centered on the cursor that are big enough in radius to appear, at least partially, in at least one of the screens. By looking at the displayed part of the circle, its position and its curvature, users can tell how far and in which direction the Perspective Cursor is located. When the cursor is barely out of one display, the displayed arc section of the halo is highly curved, showing most of the circle. If the cursor is very far away, the arc seen will resemble a straight line.

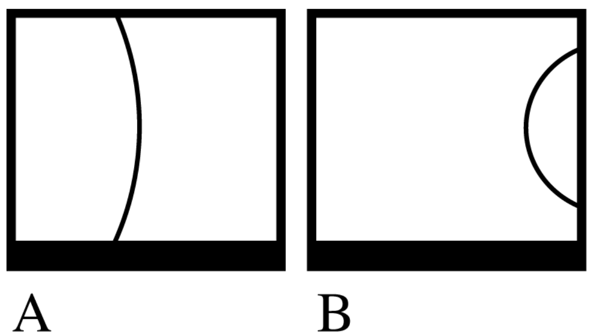

Figure 50. Two examples of halos. A) the cursor is far to the left of the screen. B) the cursor is close to the right of the screen.

The perspective variant of Halo (Perspective Halo) works on the same principle, but instead of showing circles of the screen, it shows conic curves that project a circle in the user's image plane (see Figure 51). The Halo curves are generated by a cone with its origin in the head position of the user, and its axis centered on the position of the cursor. The angle of the cone changes dynamically to make sure that the halo is shown by at least one display when the cursor is not on any display. See also accompanying video figure (Perspective-halo.wmv).

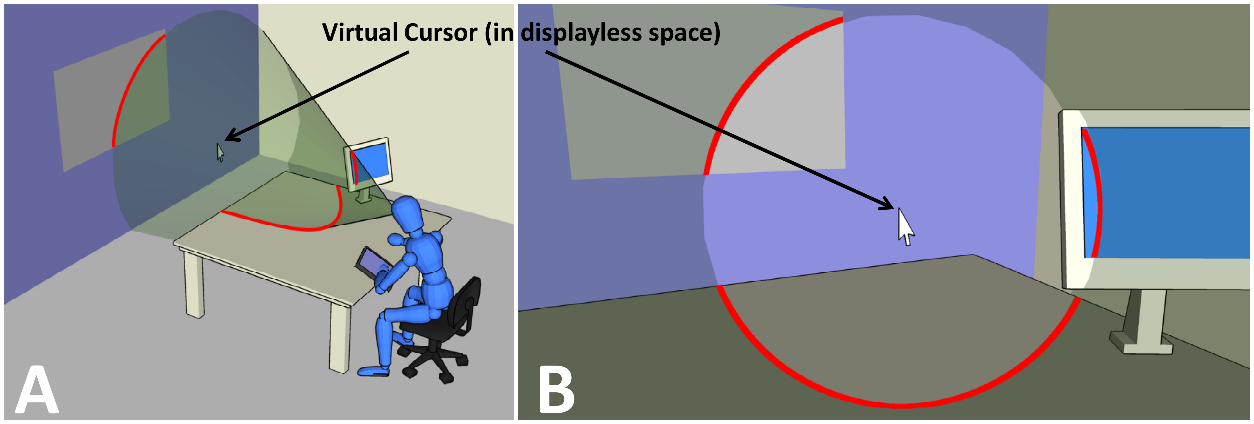

Figure 51. The geometry of Perspective Halo seen from A) an external point of view and B) the point of view of the user. The halo that is visible in the displays (red lines on the tabletop display, the monitor and the wall display) is generated by a virtual cone. Neither the virtual cone nor the virtual cursor would be visible in a real environment (they are visible here only for illustration purposes).

5.3.4. Participants

The experiment was conducted with 12 right-handed participants (2 females and 10 males) between the ages of 19 and 35. The participants were recruited from the student body of the University of Saskatchewan, and were paid $10 as compensation for their time. All participants had experience with graphical user interfaces. Participants signed a consent form and took a demographic questionnaire before the study, and filled a post-study questionnaire after. Samples of the consent form and the questionnaire are reproduced in Appendix A. Each subject was tested individually. Each experiment took approximately 70 minutes to complete.

The experiment and the handling of participants were carried out according to the University of Saskatchewan's guidelines for experiments with human subjects.

5.3.5. Task

The subjects were asked to click on an origin icon (70x70 pixels) and then again in a destination icon (of the same size) as fast as they could, but without sacrificing accuracy. Both icons were visible and seen by the user before each task started. The pairs of origin/destination icons were selected according to the results of a pilot study that provided us with four groups of tasks that represent different kinds of multi-display interactions: simple across-displays, complex across-displays, within-display and, high-distortion within-display.

5.3.5.1. Simple Cross-display Tasks

In these tasks the spatial relationship between the origin display and the destination display is simple (Figure 52.A). In our setting this is the case only for the tabletop display and the wall display (1 and 2 in Figure 47, page 83). The transition between these two displays is easy because both are more or less the same size, the connecting borders are parallel, and the displays are at a right angle.

5.3.5.2. Complex Cross-display Tasks

In this group of tasks the origin and destination displays are not aligned in any way, and they might be of very different sizes (Figure 52.B). These tasks are the most interesting from the point of view of this study because they represent the transitions of the complex scenarios where the difference between perspective and planar techniques is most evident.

5.3.5.3. Within-display Tasks

These are tasks with the two icons in the same display, i.e., single-display tasks (Figure 52.C). This group of tasks was introduced in the study as a control to test whether techniques that perform well for multi-display transitions are also good at single-display interaction.

5.3.5.4. High-distortion Within-display Tasks

In a display that is located very close to the user and in a parallel angle with the line of sight (e.g., our table top) there are regions of the display (the corners closest to the observer) that suffer from a strong perspective effect or scherzo that affects control perception. As the pilot study suggested that tasks operating in these regions would yield specific effects, we created another group testing this kind of tasks (Figure 52.D).

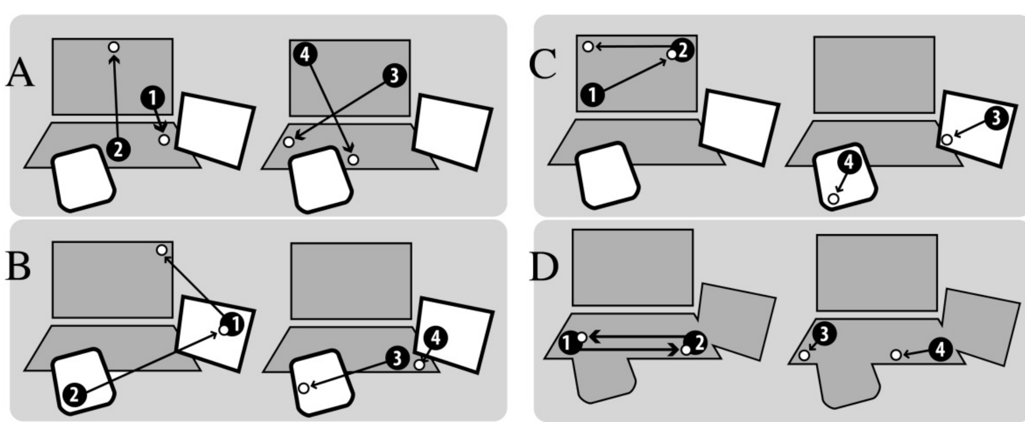

Figure 52. Task types: A) simple across-displays B) complex across-displays C) within-display D) high-distortion within-display.

The system considered a trial a miss if the second click did not fall inside the area of the destination icon, and gave distinctive auditory feedback for hits and misses. The task completion time was measured between the click on the origin icon and the click on the target object.

5.3.6. Experimental Design

The experiment used a 3x4 within-participants factorial design with planned comparisons. The factors were:

- Interaction Technique (Perspective Cursor, Laser Pointing, Stitching)
- Task type (Simple across-displays, complex across-displays, within-display, and high-distortion within-display).

The experiment comprised 3 blocks of trials, one for each interaction technique. Each block was split into two sets of trials; a training set and a test set. The training set had 32 trials, two per task, while the test set consisted of 8 trials for each of the 16 tasks in a random order, for a total of 128 test trials per technique (32 per task group). Each subject provided a total of 384 valid time measurements. The order of the conditions was balanced across subjects (2 in each possible order of interaction technique).

The number of trials was determined through a conservative a-priori power analysis (power = 0.8, estimated standard deviation = 0.62) that made our experimental design capable of detecting differences in means larger than 15%.

For each trial, completion time and hit/miss information was recorded. At the end of all trials the subjects were asked to complete a questionnaire ranking the three techniques in terms of performance, accuracy and preference. They were also asked to fill out a workload assessment form for each technique.

5.3.7. Results

Four sources of data were gathered: completion time, accuracy, user preference, and workload assessments.

5.3.7.1. Completion Time

An ANOVA test over all the successful trials with interaction technique and task type as factors, and the participant as random factor showed clear main effects of interaction technique ($F_{2,22} = 8.8$, $p = .002$) and task type ($F_{3,33} = 94.6$, $p < .001$). The statistical interaction between interaction technique and task type was also significant ($F_{6,66} = 20.5$, $p < .001$).

Figure 53 shows the average completion times and standard errors for the three techniques grouped by task type. As the general ANOVA test indicated that there was an interaction between interaction technique and task type, we proceeded to analyze the effects of interaction technique for each of the conditions of task group.

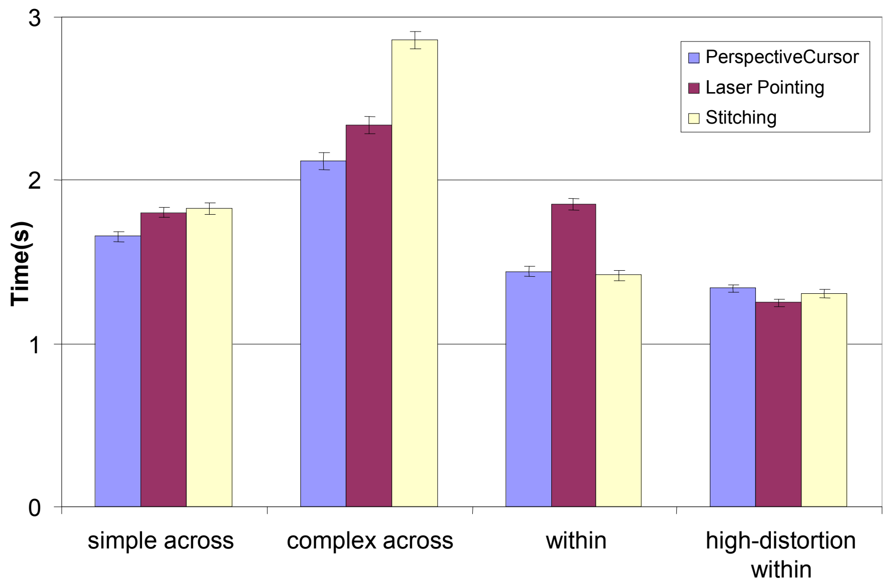

Figure 53. Time of completion in the different conditions. Error bars indicate standard error.

Table 7. Average completion times for each technique (columns) and each task type (rows - in seconds). Values between parentheses represent standard error.

	Perspective Cursor	Laser Pointing	Stitching
Simple across-displays	1.66 (0.03)	1.80 (0.03)	1.82 (0.03)
Complex across-displays	2.12 (0.05)	2.34 (0.06)	2.86 (0.05)
Within-displays	1.44 (0.03)	1.85 (0.03)	1.41 (0.03)
High-distortion within-displays	1.33(0.02)	1.25(0.02)	1.31(0.02)

Simple Across-displays Tasks

For simple transition tasks the ANOVA test with technique and task as main factors revealed that there were differences in performance amongst the techniques ($F_{2,22} = 4.01$, $p < .033$). The Tukey-HSD multiple-comparisons post-hoc test showed that for this group of tasks, Perspective Cursor ($t_{avg} = 1.6563$s) is significantly faster than Laser Pointing ($t_{avg} = 1.803$s) and also faster than Stitching, ($t_{avg} = 1.8263$) but these two are not significantly different from each other.

Complex Across-display Tasks

For complex transition tasks, the ANOVA test revealed that there were also differences in performance amongst the techniques ($F_{2,22} = 16.24$, $p < .001$). The Tukey-HSD multiple-comparisons post-hoc test showed that all task groups were significantly different from each other. Perspective Cursor was the fastest ($t_{avg} = 2.12s$), followed by Laser Pointing ($t_{avg} = 2.34s$), and Stitching ($t_{avg} = 2.86s$).

Within-display Tasks

For the tasks involving only one display, an ANOVA revealed that there were also differences in performance amongst the techniques ($F_{2,22} = 42.52$, $p < .001$). The Tukey-HSD test showed that the two fastest techniques, Stitching ($t_{avg} = 1.418s$) and Perspective Cursor ($t_{avg} = 1.44s$) were not significantly different from each other, but both were significantly faster than Laser Pointing ($t_{avg} = 1.85s$).

High-distortion Within-display Tasks

In the tasks that took place in areas of high perspective distortion, the ANOVA test did not reveal differences in performance ($F_{2,22} = 1.57$, $p = 0.23$), and therefore we did not perform post-hoc tests.

5.3.7.2. Accuracy

In terms of overall accuracy Perspective Cursor was the best interaction technique with 45 misses, followed by Stitching with 57, and Laser Pointing with 154. Figure 54 shows the distribution of missed targets over different task types. In the high-distortion-within-display tasks there were 5 misses for each interaction technique. In all other task types Laser Pointing had the most number of misses with more than 35 misses per task.

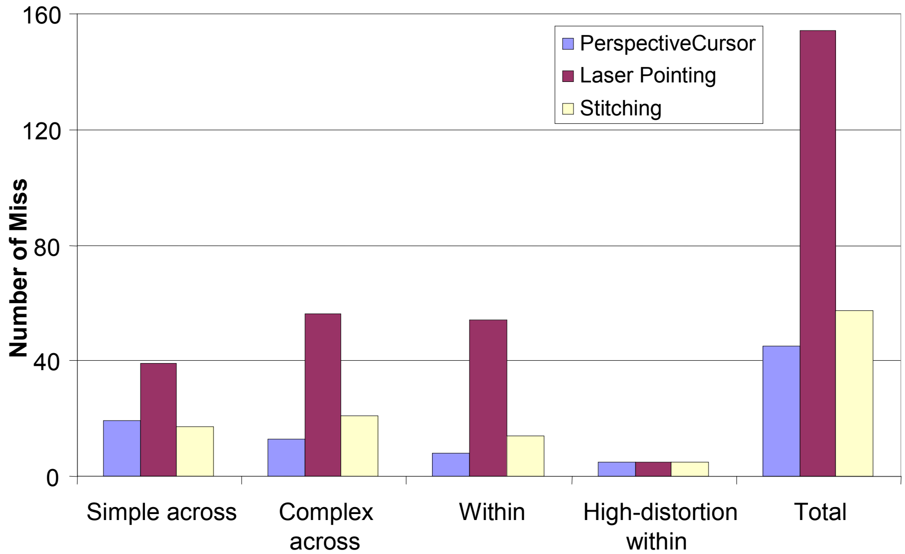

Figure 54. Number of misses per condition (of a total of 4608 trials).

Table 8. Number of misses/hits for each technique (columns) and each task type (rows).

	Perspective Cursor	Laser Pointing	Stitching
Simple across-displays	19/365	39/345	17/367
Complex across-displays	13/371	56/328	21/363
Within-displays	8/376	54/330	14/370
High-distortion within-displays	5/379	5/379	5/379

5.3.7.3. User Preference and Workload Assessment

After finishing the tasks, users were asked to rank the techniques in order of subjective speed, accuracy and preference. Most users perceived Perspective Cursor as the fastest technique (9 first places and 3 second places) over Laser Pointing (3 first, 6 second, 3 third) and Stitching (3 second, 9 third). Perspective Cursor was also considered the most accurate technique (10 first, 3 second) followed by Stitching (2 first, 5 second, 5 third), and Laser Pointing (5 second, 7 third).

When asked to rank the techniques by preference, Perspective Cursor was preferred by all but one user (who ranked it second). Laser Pointing received one first place vote, 8 second and 3 third, followed by Stitching, which received three second places and 9 thirds.

The users were also asked to fill out a workload assessment questionnaire. Analysis of the questionnaire showed that there were significant differences in user's perception of frustration, mental load and physical effort for the different techniques. Across these categories users considered Perspective Cursor less frustrating, easier to handle mentally and less physically tiring. Of particular interest is the assessment of physical effort, in which Laser Pointing received an average of 5.91 out of 7, much higher than Stitching (3.41) and Perspective Cursor (2.41).

5.3.8. Discussion

5.3.8.1. Perspective vs. Planar Techniques

The experiment found that the three techniques perform differently for across-displays tasks. Perspective Cursor is the fastest when several displays are involved (up to 26% faster than Stitching and 8% faster than Laser Pointing). Laser Pointing is better than Stitching when the relative position of the displays involved does not allow a straightforward stitching. In this kind of interaction Stitching is confusing to the users. The performance of Perspective Cursor and Laser Pointing were more similar to each other than to the Stitching technique, except for the case of single-display tasks, where the Laser Pointing was clearly inferior. These results provide evidence that in complex environments, planar mappings of the input space present problems that are solved by using perspective information. We observed that users had difficulties remembering how to access one display from another when using the Stitching technique, whereas there was no need for this with perspective techniques. Several subjects reported that they needed to plan the movements of the mouse ahead according to the stitching scheme.

Although the comparison with Laser Pointing helps further support that perspective is useful in CDOM, it is the large difference between the performances of Perspective Cursor and Stitching that is most revealing because the techniques use the same input device.

As we expected, a simple layout of monitors does not create serious problems for the Stitching technique, but Perspective Cursor is also faster than Stitching for these transitions (Perspective Cursor is 8% faster on average). One might think that the blank space that the Perspective Cursor has to cross between displays increases the interaction completion time, but consistent with what Baudisch et al. report [2004], it is the lack of that space in Stitching that makes the transition less natural and slower, as the users end up overshooting much more often than with other techniques.

It should be noted also that the 3D geometry characteristics of perspective-based techniques allowed a seamless interaction across displays of very different resolutions without an explicit change in C/D ratio.

5.3.8.2. Perspective Cursor

In all but the high-distortion tasks Perspective Cursor was the best technique, or at least not significantly worse than the best. In addition, Perspective Cursor was as fast as Stitching in the simple within-display tasks, which means that the multi-display capabilities of the technique are not traded off for a poorer performance in the standard single-display interactions that we are used to.

The overall results for Perspective Cursor show that there is value in using a relative control device like the mouse in combination with perspective. Users seem also to appreciate it, as all but one ranked it best. We think that Perspective Cursor, although relatively complicated to implement compared to a planar technique, is a more efficient option for control of multi-display environments than the existing alternatives. If efficiency is an important element for the design of a particular MDE, the gains in preference and performance will probably justify the use of position tracking technologies.

One possible problem of Perspective Cursor is the possibility of losing the cursor in non-displayable space. We were worried that users could have trouble with this feature but the halos seemed to work well, and when the users lost the cursor, they were capable of bringing it back to a display very quickly by looking at the halos.

It must also be mentioned that the relative-positioning nature of Perspective Cursor might allow for further target-acquisition optimization as in [Grossman and Balakrishnan, 2005; Balakrishnan, 2005].

5.3.8.3. Laser Pointing

For tasks that involved complex display transitions Laser Pointing proved to be of value (almost 20% faster than Stitching). The experiment also provided weak evidence that in situations of high perspective distortion (closest corners of a non-perpendicular display) this technique is preferable to mouse-based techniques.

However, the accuracy of Laser Pointing was far below the other two (89% of success compared to 96% of Perspective Cursor and Stitching). This problem is due to the inherent inaccuracy of the device and has been reported many times [Myers et al., 2002; Olsen and Nielsen, 2001]. Although there are ways to improve this accuracy by filtering or changing the interaction techniques [Davis and Chen, 2002; Parker et al., 2005; Myers et al., 2001], our main focus for this experiment was on performance and so we decided to include the technique without modifications that may introduce feedback lags or arbitrary delays.

Laser Pointing is faster than Stitching for complex transitions, but slower for single-display movement. This suggests that, even though Laser Pointing is inferior to Perspective Cursor (likely because of the physical characteristics of the input device), it still retains some of the advantages of being a perspective interaction technique. Some of the difference in performance between Laser Pointing and Perspective Cursor might be explained through the parallax between the pen and the point of view of the user, which is typical of Laser Pointing usage, but is absent from Perspective Cursor. The design of our study makes it impossible to confirm or support this hypothesis with the data and further studies will need to be designed to investigate this.

Another issue of laser pointing techniques is how we provide a button click. If the button is in the device itself, accuracy is further decreased by the clicking movement at the moment when most stability is required: at target acquisition. This problem can be solved if we perform the clicking gesture using the non-dominant hand, but this raises other kinds of problems for real-life environments because we usually need the non-dominant hand for other purposes (e.g., holding another device, gesturing, etc.).

Another drawback of Laser Pointing made evident by the data collected is that it is a very tiring technique. Several users reported this, and the technique was rated the most physically demanding. Although we agree that in a real-life situation the use of a laser pointer or a pen would not be as intensive as in our experiment, this effect should be considered for applications that require intensive pointing for long periods of time.

It must also be mentioned that the technological limitation that reduced accuracy when the pen was too close to the tablet might have had an effect on the trials that involved the tablet.

5.3.8.4. What You See Is What You Can Control

One important aspect of Perspective Cursor is that it provides control only over the display surfaces that are visible, and in the degree in which they are visible. This means that Perspective Cursor may not be adequate for environments in which the multi-display interaction is intended for full resolution control of non-visible displays or machines from a single interface. For these situations it would be better to use remote-control techniques like Mighty Mouse [Booth et al., 2002].

In general, perspective techniques are limited to the control of display surfaces that are visible from the point of view of the user. For example, Ashdown and colleagues' CDOM technique based on head tracking [Ashdown et al., 2005] or Dickie and colleagues' gaze detection system to activate devices [Dickie et al., 2006] would not be able to specify displays that are not visible to the user. This is not necessarily a disadvantage of perspective techniques; most of the time it does not make sense to try to manipulate objects that are not visible (i.e., when the feedback loop is broken - I discuss this issue in detail in the next chapter). For scenarios in which it might be desirable to put an object in a display that is not visible (e.g., if we want to send a document to a laptop that is already stored in a backpack, or to a PDA that is inside a pocket), techniques that are not spatial are worth considering (see Chapter 4); if an arbitrary location in the input configuration is allocated to displays that are not visible (or not present) in the immediate space, we run the risk of cluttering the space and making interaction unnecessarily complicated and unintelligible.

5.3.8.5. Implications for Collaborative Environments

Perspective techniques might be difficult to interpret and predict for users that are not in control of the cursor or the input device. This could pose problems in collaborative environments where it might be important to constantly acquire accurate workspace awareness of what others are doing. This problem is likely to be more acute for Perspective Cursor than for Laser Pointing techniques because the indirect nature of the mouse movement and the small area in which mice are typically used are less visible to collaborators.

Planar techniques are not exempt of this problem (Stitching's gesture-cursor movements relationships can also be difficult to interpret by others); however, 3D relationships between a user and the environment are likely to be easier to interpret than planar relationships. In terms of visibility and interpretability of the gestures, literal techniques are superior because the relationship between the manipulated object and the user that is doing it is straightforward and clearly visible. It might be possible to add embodiments to indirect techniques from the perspective and the planar groups so that the disconnect between action and gesture is reduced. Although we have explored this topic for single-display groupware [Nacenta et al. 2007; Pinelle et al. 2008; Pinelle et al. 2009], future research adapted to the particular issues of MDEs is needed.

5.3.8.6. Applicability and Costs

Perspective techniques require some kind of information about the point of view of the user; for example, Perspective Cursor requires tracking of the user's head relative to the position and orientation of any display in the room. Although current alternatives for the implementation of this tracking are still expensive (e.g., 3D magnetic trackers, computer vision tracking, or active sensors), the price of these technologies has been reduced dramatically in the last few years and will likely keep falling. Each of these technologies still has their own problems (e.g., tethered sensors, interference from metallic objects, line of sight occlusions), and the tradeoffs between different technologies or not using any tracking technology at all (manual configuration) should be carefully considered by designers.

In contrast, some planar techniques might not require any special sensing of the environment or the position of the user. MDE designers should seriously consider planar techniques against perspective techniques if performance is relatively unimportant or if the MDE is simple enough that a planar configuration will still be a good approximation. It should be noted, however, that even simple planar techniques as Radar might pose important problems for interaction if there is no knowledge about the position of the user (WIM views might appear rotated to certain users, which causes difficulties in its operation [Kortuem et al. 2005]). This problem is also evident for indirect systems (e.g., mouse controlled) that make use of horizontal surfaces; the location of the users should be taken into account to provide an effective input model to physical configuration mapping (see Figure 55).

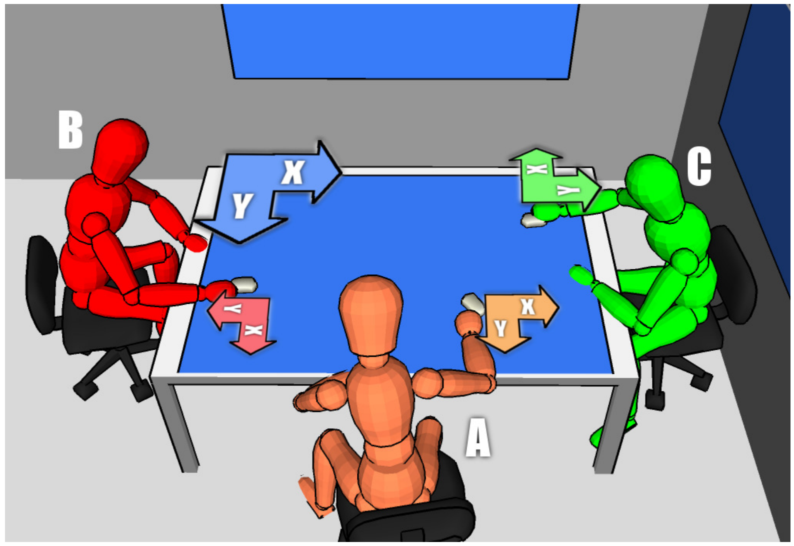

Figure 55. In an environment with horizontal displays that is controlled by indirect devices (e.g., mice), the position of the user directly influences the mapping of the input space and the physical configuration (the X and Y displacements of the mouse have different effect depending on where the user is with respect to the table).

5.3.8.7. Generalizability of the Results

This study compared different techniques for multi-display pointing. The experimental setup included displays of several sizes in several positions, which made the controlled environment reasonably similar to real conditions; however, there is still much to learn about these techniques in more general situations, in particular: how do the techniques perform in tasks other than pointing (e.g., text selection, drawing)? Are these techniques equally useful in multi-user environments? Can the techniques be adapted to other kinds of input devices? These questions have to be answered through future experiments.

5.4. Chapter Discussion

This section looks at the two experiments of this chapter from a more general perspective and provides a more general discussion of what they tell us in the context of this dissertation.

Experiment 3 tested a number of planar and literal interaction techniques. The results indicate that literal techniques have an important performance advantage over planar techniques. This advantage is consistent with the Dimensional Overlap model introduced in Sections 3.1.4.2 and 4.3; a perfect overlap of the input model with the physical configuration of the space seems to be a better adaptation to the nature of the human perceptual-motor system. Note that these results contradict a naïve (but extended) interpretation of Fitts's Law for interface design in which a larger gesture is assumed to imply poorer performance than a shorter gesture. The results of our study show numerous cases in which larger gestures resulted in performance improvements (e.g., the Pick-and-Drop vs. Pantograph comparison). The correspondence between input and output must, therefore, be taken into account when choosing interaction techniques and input devices since our study shows that it can be the determining factor.

Although literal techniques exhibit good performance, this does not automatically make them the group of choice; these techniques suffer a very important limitation when the MDE is a large space (and having to move to touch objects becomes tedious or even impossible – e.g., for very tall displays), or when the space is collaborative and we cannot afford crossing the personal and work spaces of others freely. This problem is an instance of a tradeoff between power and performance (or 'naturalness') that is often found in interaction design; we can design interactive systems that try to replicate as closely as possible the behavior and characteristics of real objects, resulting in interactive systems that are 'natural' and where we can leverage the knowledge that users already have about their daily environments (for an example, see [Agarawala and Balakrishnan, 2006]). These systems will, however, be largely limited by the same constraints of the real world, and lack the "super powers" that computer-based systems are capable of (e.g., reaching beyond hand's reach with the Pantograph). On the contrary, if we design to take full advantage of the abstract capabilities of computerized electronic systems we might end up designing interfaces that are difficult to understand by users (techniques might be perceived as 'magic') and that ignore good human factors design, which might have negative repercussions on crucial elements of the workflow of users with the system. For example, long reach without a proper environment might disturb territoriality [Scott et al., 2004; Pinelle et al., 2009].

Among the interaction techniques that were studied in Experiment 3, Radar is perhaps the most peculiar and the most difficult to classify. The task design for Experiment 3 allowed Radar to interact with the small objects in a literal fashion, without really requiring looking at the large displays that compose the real MDE. This is not realistic in a real environment because operation and visualization of data in the multiple and large displays is usually the reason why the MDE is used in the first place. In other words, what is the point of building an MDE with wall displays, tabletops, monitors and other devices if it is only used through a small interactive display in front of the user? In more realistic scenarios, we are likely to observe significant reductions in performance due to the extra mental processing required to match elements and spaces in the miniature and the real space. This opens a new set of research questions and possible interaction technique designs that fall out of the scope of this dissertation, but that could be important for future MDEs.

Radar or, more generally, world-in-miniature techniques (WIM) are classified in this dissertation as literal or planar (depending on whether their operation can take place completely within the miniature or not). However, it is also possible to apply perspective to the representation of the radar (a technique that has not been proposed yet), and there might be situations in which operating exclusively through the miniature might be beneficial (e.g., in situations similar to that portrayed in Figure 4.E of page 5). World-in-miniature approaches probably deserve a separate sub-category, or even a separate level of analysis; however, this dissertation does not include it because it is not a specific issue of MDEs. For a classification of techniques that accounts for the specificities of world-in-miniature techniques in single-display groupware see [Nacenta et al., 2007].

Experiment 4 was aimed at comparing planar and perspective techniques. Findings from this experiment indicate that, consistent with the predictions of the Dimensional Overlap model, performance is significantly affected by the overlap between the input model and the physical configuration. This is seen clearly in the performance advantage of the two perspective techniques over the planar technique in tasks where the planar representation deviates most from the physical environment.

Both experiments also show how the input device and the particular design of a technique is also an important factor to explain differences between techniques (besides the type of mapping): Perspective Cursor and Laser Pointing use both perspective mappings, but Laser Pointing shows accuracy problems and less performance; Pick-and-Drop, Synchronized Gestures and Radar are all planar techniques, but there are large differences in performance stemming from the differences in the amplitude and shape of the gesture.

In general, the two studies of this chapter indicate a general tradeoff between performance, power, fitness for collaborative environments, and cost. Literal techniques are the fastest, and provide good workspace awareness to other in the workgroup, but their effective range is limited to the user's direct reach (lack of power). Planar and perspective techniques do not have the reach limitation, but are poorer in performance and in the amount of information that they convey to others. Perspective techniques might improve the performance of planar techniques, but they are (currently) more costly to implement and, perhaps, more difficult to interpret by others.

This does not necessarily mean that the choice of interaction technique does not matter, or that designing new interaction techniques is not important. The results of our study show that the design of a technique is not a zero-sum game; for example, perspective techniques are a superior choice to planar as long as the alignment of the displays is not trivial (which makes them practically equivalent). Ultimately, the onus is on the designer of the MDE, who has to make a balanced choice and match between techniques according to the specific needs of the designed environment and task. This dissertation should, however, facilitate this decision by providing empirical data upon which to base the design decisions.

It is also important to remark that these two experiments are not a fully comprehensive exploration of the space, even in this chapter. First, the number of techniques that can be tested is necessarily limited by the restrictions of experimental design (there are practical limits in the number of conditions – techniques – that can be tested in an experiment). It is possible that some of the untested techniques present anomalies in terms of performance.

Second, the focus of this dissertation is on performance and user preference. This approach should be complemented with other studies focused on collaborative aspects and on emergent behavior in more open scenarios.

Third, there are many questions still open for study. I justified the lack of a comparison between literal and perspective techniques at the beginning of Section 5.3 (page 82), as a less interesting case (literal techniques are limited in range), and one that has found previous partial solutions in literature. I assume that the speed that can be achieved through literal techniques is difficult to beat, even with perspective techniques (which are faster than the planar techniques that were tested against it). However, further research might uncover ways to make indirect techniques faster than literal techniques, if not in all situations, perhaps in certain scenarios or for certain types of configurations. This topic presents an interesting avenue for future research. Similarly, I have discussed already the special nature of world-in-miniature techniques, and how their performance cannot be fully compared yet with other technique types.

Nevertheless, these two studies, together with the framing of the problem of mapping between input model and physical configuration, and the identification of the three groups of techniques (planar, perspective and literal), provide a common language for researchers, a solid foundation for future research, and a source of empirical data that can help designers build better MDE interfaces.

5.5. Implications

There are five main lessons from this chapter that can translate into advice for the design of CDOM interfaces.

- Literal techniques are fast, but are limited to the user's physical reach. Choose literal techniques whenever most interaction happens within hand's reach of the user.
- Perspective techniques exhibit better performance than planar techniques in scenarios with complex display setups (such as the one used for experiment 4), but are also more difficult to implement and require sensing technology that might increase the cost of the system. If the MDE is complex (e.g., if displays are of many different sizes, if the MDE includes mobile displays, or if the positions of displays are necessarily irregular due to the physical constraints of the space where it is deployed) and performance is important, implementing a perspective technique might be worth the extra cost in sensing technology.
- Although the type of mapping (planar, perspective or literal) seems to be a stronger predictor of performance than the input device or other details of the technique, these can still help explain part of the variance between techniques. In particular:
    - Choose forward or aligned mappings over backward mappings (Pantograph vs. Slingshot)
    - Mapping non-spatial variables to spatial interaction may hamper performance (Press-and-Flick)
    - Mice are preferable for pointing tasks to remote pointing devices (Laser Pointing) because of their superior performance and increased accuracy, although collaborative concerns against the interpretability of gestures performed with mice should be taken into account.
- When no empirical data is available, trying to achieve high dimensional overlap between the gesture and the environment is a good idea.

- User preference seems to follow performance results, at least for tasks where performance is required. If no other concerns are present (collaborative, physical tiredness, etc.) aim at making the interaction faster.

5.6. Conclusions

In this chapter I have identified the mapping between the input model and the physical configuration of the MDE as an important design factor that significantly affects performance for spatial CDOM interaction techniques. I distinguish three categories of techniques according to the mapping: planar, perspective, and literal.

Two experiments were designed and executed to investigate the differences between the different groups of techniques. Experiment 3 shows that literal techniques perform faster than planar techniques, although they are inherently limited by the short range. Experiment 4 compared planar and perspective techniques in a series of tasks with different relationships between displays. This experiment found that perspective techniques, as predicted by the Dimensional Overlap model, provide a performance advantage over planar techniques.

Besides the differences between technique types, the experiments also revealed a number of differences due to the implementation and input devices of different techniques.

CHAPTER 6: MOVEMENT EXECUTION

The three previous chapters have dealt with parts of the process of cross-display object movement that are crucial for the planning of the CDOM action. In Chapter 3, we looked at how different ways of referring to the destination display – mostly a mental process – affect CDOM. In Chapter 4 and 5, we looked at how the relationship between the physical configuration of the MDE and its representation in the system (the input model) affects the speed and accuracy with which object movement can be achieved. This chapter deals with the actual execution of the action.

Execution of the cross-display action is altered in MDEs with respect to single-display environments because the display space is not continuous; in other words, the display space is fractured and there is *displayless space*[13] in between the display surfaces. The existence of this space affects how feedback is provided and how consistent the visual feedback is with respect to the motor space, resulting in different design decisions for interaction techniques.

This chapter analyzes several aspects of interaction techniques affected by displayless space: the presence of a *closed loop* between the system and the user (as defined by Card and colleagues [1983]), the consistency between the motor and the visual spaces, the spatial warp of objects between displays, and the effect of extra motor space. The presence of feedback is used at a higher level to classify techniques into three groups: open-loop, intermittent, and closed-loop techniques. I then explore in more detail the tradeoffs in the design of indirect input techniques with respect to displayless space (Experiment 5).

6.1. Background

This section provides a review of related work, starting with the more general area of motor control, and then narrowing down the subject to issues specific to CDOM interaction techniques. The section also includes the classification of existing interaction techniques into three groups: closed-loop, open-loop, and intermittent.

[13] Displayless space is the area around the user that is not covered by a display capable of showing electronic information.

6.1.1. Studies of Performance in Motor Movement (Closed Loop)

In 1899, Woodworth performed a series of experiments to investigate the speed-accuracy tradeoff in aiming movements [Woodworth, 1899]. He was the first to identify the speed-accuracy tradeoff of targeted motor movements and the first to measure the importance of visual feedback in accuracy. He also proposed that aiming movements are comprised of two phases: the initial impulse towards the target, and a deceleration phase to home in on the target.

More than half a century later, Paul Fitts formulated what is now probably the best known model in HCI: Fitts's Law. Although the initial formulation was derived from repetitive side-by-side tapping on vertical stripes, Fitts's Law has been extended to single targeting movements [Fitts and Peterson, 1964]. Since then, a large amount of work has used, analyzed, extended, and derived guidelines from Fitts's original work (e.g., [MacKenzie, 1991; MacKenzie et al, 2001; MacKenzie and Buxton, 1992; Soukoreff and MacKenzie, 2004]). Importantly, Fitts's Law has been shown to work for tasks and devices that do not require direct input (the original pen from Fitts's experiment) but are, instead, indirect, or even relative (the use of a mouse or a trackball for targeting on a screen). The formula for Fitts's Law generally indicates that the time to complete a targeting movement depends on the logarithm of the distance to the target *(D)* divided by the size of the target (*W*) (see equation 1). In other words, an increase in the distance to the target or a decrease of the width of the target will reduce the targeting performance.

$$(e1) \qquad Targeting\ Time = a + b \log\left(1 + \frac{D}{W}\right)$$

The main strength of Fitts's Law is that it contains within one formula[14] the variables that relate to both performance and accuracy. This allows us to compare user and device performance without having to resort to multiple comparisons (e.g., one for each of the possible constraints of accuracy). In other words, Fitts's Law allows us to find a constant for each device and person that describes the bandwidth of the human-machine system, regardless of the characteristics of the pointing task (distance) or the accuracy required (i.e., width of the target).

Fitts's Law is an invaluable tool for research because it describes and quantifies the process of targeting; however, the model does not offer any explanations as to the inner workings or the

[14] There are many alternative formulations of Fitts's Law from the HCI and Psychomotor literature, but this formulation is simple and serves our discussion well.

targeting process (it just models it). There are various underlying models of motor control for aiming movements; the most successful model to date corresponds with Woodworth's two-phase model of aiming. The Optimized Initial Impulse Model [Meyer et al., 1988] suggests that an initial ballistic movement is made towards the target, and if successful, then the task is accomplished. If not, a secondary movement is undertaken and this process repeats until the target is acquired. To optimize aiming movements under the constraints of the speed-accuracy tradeoff in this model, most movements consist of a high-velocity initial phase (ballistic phase), and a series of slower, visually-guided, feedback-corrected movements (homing-in phase). Prior to the Optimized Initial Impulse Model, other motor control models emphasized the ballistic phase (Impulse Variability model [Schmidt et al., 1979]) or the homing-in phase (Iterative Corrections model [Crossman and Goodeve, 1983]).

Most of the models and empirical evaluations done in this area are concerned with closed-loop scenarios (with a few exceptions such as [Plamondon, 1995]). In these situations a visual feedback or feedforward[15] channel gives the user up to date information that allows the correction of the gesture while it is still in progress.

6.1.2. Open Loop

If there is no feedback channel that allows the user to correct the gesture, the execution of the gesture is *open-loop* (see Figure 56). There are three possible situations in which there is no loop: when there is lack of feedback, when the gesture has to be so fast that the feedback cannot be used, and when the control of the movement is removed before the action is finished.

[15] In this dissertation I do not enter into the discussion of the distinction between feedback and feedforward that is made in the motor control literature. The only element of interest here is whether there is a channel that allows the correction of the movement, and whether this channel can be effectively used. It is mostly irrelevant for our purposes whether the information is used towards a correction of a previous part of the movement or towards the planning of an adapted subsequent sub-gesture.

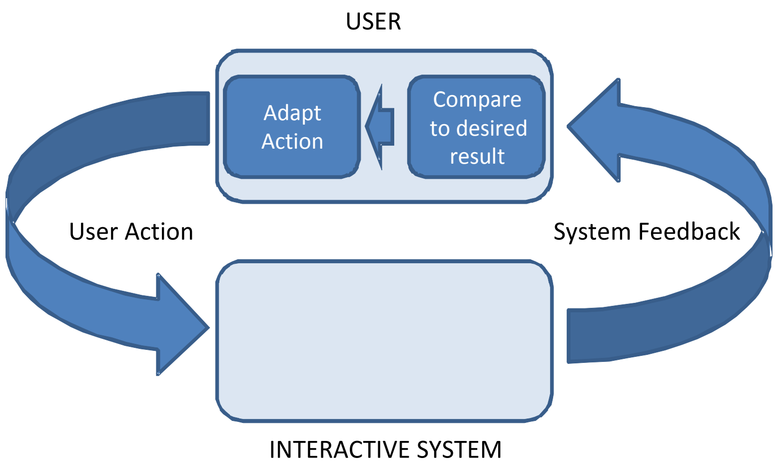

Figure 56. A simplified representation of the human-system feedback loop.

6.1.2.1. Lack of Feedback

Lack of feedback is the most common way to open the loop, and will also receive the most attention in this chapter. However, the loop is only rarely absolutely open: in very few instances is feedback completely absent. Even if visual feedback is lacking, in certain circumstances (e.g., when navigating *displayless* space with an indirect or relative input device), there are other sources of information available to the user such as proprioception (the ability to sense the position of one's own body parts). Nevertheless, visual feedback seems to be considered the most relevant and the one accountable for the largest part of the variability (e.g., [Meyer et al., 1988]).

Although there is still some controversy within the human motor performance literature as to what is the role of feedback in the performance of aimed movements, most authors agree that rapid reaching motions are adapted using visual feedback and that the lack of feedback results in a sharp decrease in performance [Elliot et al, 1999; Meyer et al., 1988]. This is true for accuracy constrained situations (when performers are asked to be accurate rather than fast) as well as for time constrained situations (when performers are forced to complete the movement within a certain time frame), except if the time constraint falls below the limit at which the feedback information cannot be processed in time anymore (this last case is the main topic of the next subsection). The lack of feedback is known to negatively affect the accuracy (number of errors – e.g., [Meyer et al., 1988]) or the completion time of the movement (e.g., [Elliot et al., 1999]).

6.1.2.2. Fast Movements

If the movement is constrained in time below a certain limit, the human psychomotor system cannot successfully integrate the feedback information into the movement performance. Performance and accuracy in these situations becomes then almost equivalent to that of open-loop situations.

The exact minimum interval required for feedback to be useful might depend on the kind of task, the individual, and the degree of expertise of the performer. As an approximation, Card et al. estimate the motor processing cycle to take between 300 and 700ms [Card et al., 1983], whereas Carlton estimates the minimum visual processing time in the order of 100-135ms [Carlton, 1992].

6.1.2.3. Interrupted Control

The last variant of open loop control has to do with the removal of the left side of the loop (as represented in Figure 56): the interruption of control. This type of open loop motor control is usual in sports. For example, when bowling or throwing a basketball, our perception of the ball going in the wrong direction or with the right weight comes usually far after we have lost control of it. Although a certain degree of control is possible during physical contact, the error in the throw is only clear later in the trajectory, and better performance is usually only possible through learning from one trial to the next. With this type of open loop, performance will be equivalent to the lack of feedback loop of the fast movements, and will only be better to the degree that the gesture movement allows for some control of the results during the controlled phase (i.e., to the extent that the movement is closed-loop).

6.1.3. Relevance to Multi-display Environments

The discussion above applies to any type of object movement, regardless of the environment where it takes place (MDE or single-display systems). However, the execution of movements is particularly relevant for MDEs because displays in most of these environments are separated by stretches of displayless space (see Figure 57).

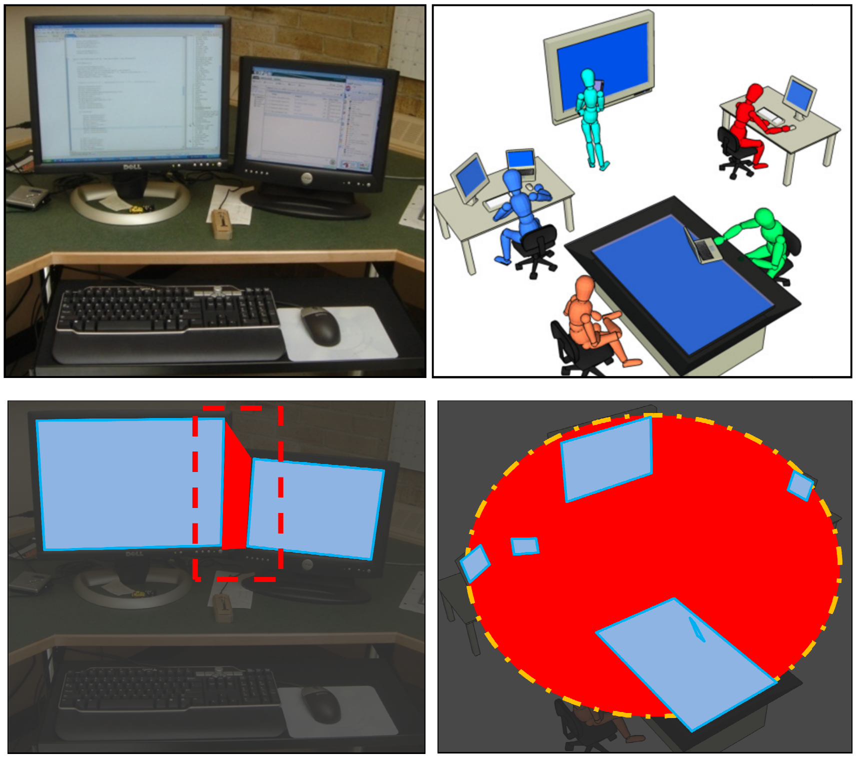

Figure 57 . Two multi-display environments (top left and top right) and their respective displayless space (bottom left and bottom right).

The presence of displayless space implies that continuous spatial interaction techniques must implement one of the following options: ignoring displayless space and provide discontinuous feedback (warping), have regions where there is no feedback, provide feedback in displayless space in a different manner, or implement a completely open-loop movement from the beginning.

Although existing literature (particularly from the Human Movement field) can provide valuable insights to predict how different techniques can affect performance, the study of movement execution in multi-display environments has received very little attention so far.

6.1.4. Types of Control Paradigm

In this chapter I classify techniques into one of three possible control paradigms: closed-loop, open-loop, or intermittent.

6.1.4.1. Closed-loop Techniques

A technique is closed-loop when there is some mechanism that allows the user to adjust the execution of the action before it is finished. This adjustment depends on feedback: for example, a pointing task in which the user can see the cursor as it moves towards the target is closed-loop, because the image of the cursor provides continuous feedback about position. Closed-loop pointing techniques have been widely studied in HCI, and conform quite closely to predictions based on Fitts's Law.

We are interested in the feedback loop as it applies to cross-display object movement; however, it is often difficult to separate single-display pointing mechanisms from cross-display object movement because techniques that allow remote placing of objects[16] also have a component of single-display-space pointing.

Techniques that have literal display configurations (Section 4.1.3) are always closed-loop because of the way we move objects in the real world: when we move an object we usually hold it all the way until it is at its destination. During the movement, we get feedback by looking at it (or our hand) and by feeling its position through the senses of touch and proprioception.

Some techniques with planar and perspective input models are also closed-loop. For example, world-in-miniature techniques allow the transfer of objects from one screen to another in a continuous mode that resembles single-display operation. The regular multi-display cursor present in current operating systems (Stitching) is also closed-loop because, although the cursor jumps in space from one display to another, there is feedback of its position at all times. Laser pointing techniques that are implemented using an actual laser are also closed-loop, since the laser spot is always visible even if it is not projected on an active surface.

Closed-loop execution is generally far more accurate than open-loop, unless the time constraints on the execution force users to eliminate the homing-in section of the gesture (see the discussion in Section 6.1.2.2). In exchange, closed-loop techniques are slower (see, for example, the comparison presented in [Reetz et al., 2007].

6.1.4.2. Open-loop Techniques

A technique is open-loop when it lacks a feedback channel or when the user cannot correct her actions before the object is in its final position (i.e., when the control loop is broken or when the gesture is too short for the control loop to make a difference).

Examples of open-loop techniques include Flick [Wu and Balakrishnan, 2003] (although not SuperFlick [Reetz et al., 2006] which closes the loop for the last part of the interaction) and the button, key, and head-tracked versions of the Multi-Monitor Mouse [Benko and Feiner, 2005].

[16]the ability of placing objects in a specific part of a display as opposed to on a predefined location of the display.

Researchers have also tried to take advantage of the increased speed of open-loop style interaction to improve the performance of tasks that are traditionally closed-loop. For example, with the Delphian Desktop [Asano et al., 2005] the cursor movement is analyzed during the ballistic phase of the targeting movement in order to determine the likely target of the action. If a likely destination is found, the cursor is warped to it in order to reduce or eliminate the homing-in time. Although it is not clear yet whether this kind of approach can leverage performance gains in real use scenarios, further research in the space between open-loop and closed-loop execution paradigms might result in improved performance.

6.1.4.3. Intermittent Techniques

Intermittent techniques arise specifically in MDEs because of the existence of displayless space between displays. If the displayless space is simply ignored (for example, in standard cursor movement), the technique is closed-loop because the cursor is always visible in at least one display. However, this causes inconsistencies between the motor and the visual feedback spaces. For example, in the regular stitching used by current operating systems, the movement of the cursor becomes irregular because it travels very fast exactly at the transition between the two displays (see Figure 58).

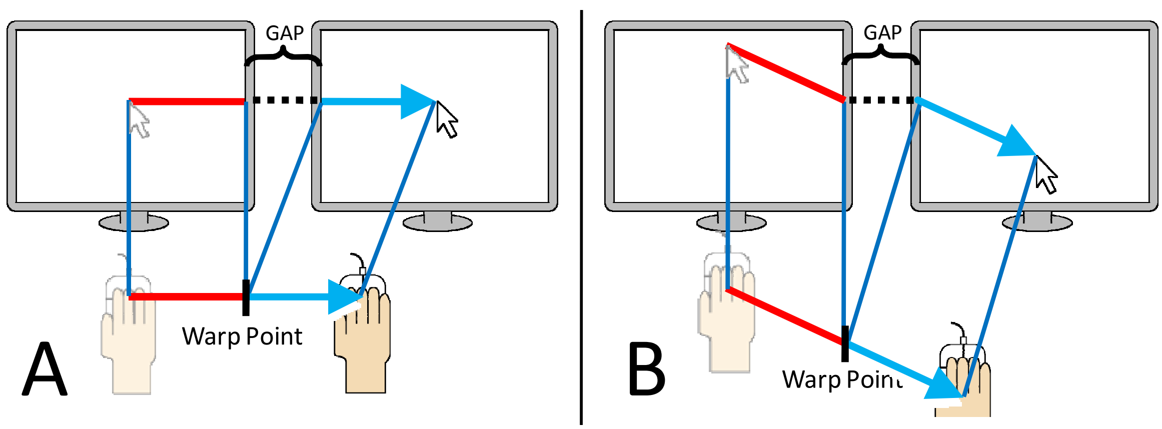

Figure 58. Inconsistency between motor space and visual feedback with Stitching. A) The gap (displayless space) is ignored in motor space (gap is compressed into the warp point). B) A diagonal motion is transformed into a multi-linear trajectory.

The alternative is a motor space that takes into account the space between the screens. This technique, first developed by Baudisch and colleagues [Baudisch et al., 2004] is called Mouse Ether. The technique introduces extra motor space so that the cursor has to travel across the displayless space before it shows up in the next monitor. The cursor is, by definition, not shown in displayless space, resulting in a CDOM movement that starts as a closed-loop action, then becomes open-loop and, once in the destination display, is closed-loop again. Mouse Ether also requires adding extra motor space to the amplitude of the movement (except if the target falls between the displays, which is rarely the case), altering the relationship between the *D* and *W* components of Fitts's Law, and therefore theoretically increasing targeting time. In exchange, the inconsistency between motor space and visual feedback shown in Figure 58 disappears.

Mouse Ether has been compared at least to one closed-loop technique (Stitching) in its initial evaluation by Baudisch and colleagues [2004]; however, their experiment compared Mouse Ether to a version of Stitching with different control-display gains[17] for each monitor. This causes a set of extra problems for Stitching that are not directly related to the specific problem of displayless space (and can be fixed simply by equalizing C-D gains in both displays). For example, the cursor warps vertically (as well as horizontally) for displays that are horizontally aligned. Moreover, the magnitude of the vertical jump is different depending of the height of the point where the crossing takes place (see Figure 59).

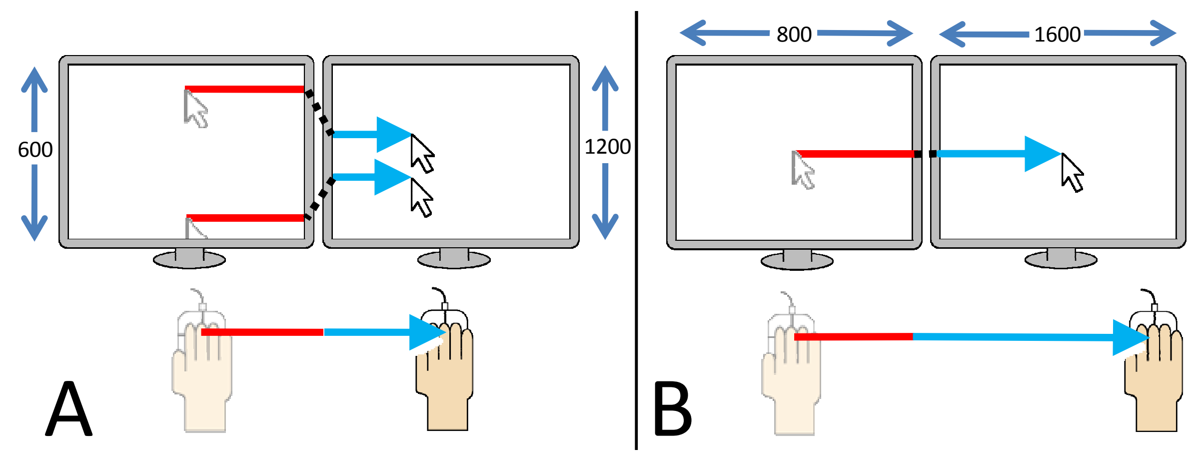

Figure 59. Motor-visual inconsistency due to differences of C-D gain between the displays. A) Alignment of all trajectories is impossible, B) the magnitude of the required movement depends on the C-D gain of the display.

Baudisch et al's experiment is, therefore, not useful to discern whether the improvement in performance is provided by the addition of the extra motor space or by the equalization of motor-visual resolution across displays. It is also not possible to detect whether the lack of feedback in

[17] The control-display gain (C-D gain) is the relationship between the movement of the cursor on the display and the movement of the input device (usually the mouse), both measured in the same units. In the literature, the term *control to display ratio* (C/D ratio) is used instead, which is the arithmetic inverse of the C-D gain.

the cross-display transition has a negative effect on performance. Unfortunately, existing work in HCI or psychomotor literature is not applicable in this context either; HCI research that considers the non-linearity of control-display relationships (e.g., [Casiez and Balakrishnan, 2007]) does not deal with the lack of feedback, and the human motor studies that deal with constrained feedback (see [Carlton, 1992]), do not apply to non-linearity in control.

6.1.4.4. Off-screen Feedback

The lack of display pixels on certain parts of the MDE (displayless space) does not completely prevent the system from providing feedback to the user about these areas. The alternative to have an open loop in displayless space is to use off-screen feedback techniques that indirectly show the position of off-screen objects.

Several techniques have been proposed for showing the presence, direction, and distance of off-screen objects. The simplest technique, used often in video games, is an arrow on the boundary of the screen that points to the off-screen object. Additional cues can be added to the arrow to convey more than just direction; this is seen in the City Lights technique [Zellweger et al., 2003], which represents off-screen objects with blocks on the edge of the display. The size, shape, and color of the blocks indicate different properties of the object. Arrows and City Lights, however, have difficulty showing the distance to off-screen objects because the mapping between visual variables (e.g., color or length of the arrow) and distance are arbitrary and difficult to recognize and to learn.

Halo [Baudisch and Rosenholtz, 2003] was designed to address this problem, by explicitly representing both distance and direction. Halo shows an arc on the edge of the view for each off-screen object; the arc is centered on the object, so the radius of the arc indicates the object's distance and the location on the display indicates direction. Halo was included as help to locate the cursor in the Perspective Cursor technique (see Section 5.3.3.3).

By using these techniques we can close the feedback loop in displayless space; however, it is not clear if the change in the type of feedback (direct or indirect) also affects performance negatively (users have to switch between direct and indirect representations of the object).

6.2. Research Questions

As we have seen in the previous sections, open-loop and closed-loop techniques have been shown to differ significantly in performance: open-loop techniques are fast but inaccurate, while closed-loop techniques are much more accurate, but slower. This issue is generally undisputed and relatively unaffected by the specific characteristics of MDEs.

Instead of focusing on the well-known general differences between open- and closed-loop techniques, I decided to investigate in more detail specific MDE CDOM execution issues that, although critical for many MDEs, have not been studied before. Experiment 5 is focused on the analysis of design decisions that affect CDOM techniques that use indirect input. In particular, the experiment looks at different ways to deal with displayless space.

Although the setup of the study still compares closed-loop and intermittent techniques, the issues involved go beyond this classification. Several issues were of interest in this study. Some CDOM display techniques (such as Stitching) provide full feedback and can be considered closed-loop techniques. In addition, because Stitching warps the cursor across displays, it also minimizes the targeting distance, which should have a substantial effect on performance. However, Stitching also fractures the display space, and little is known about how these fractures affect CDOM. Similarly, techniques that explicitly include displayless space in the virtual workspace (such as Mouse Ether) are intermittent techniques, but can be augmented with varying forms of synthetic feedback. The experiment was designed to investigate several specific questions within this set of issues.

The experiment in this chapter analyzes a narrower aspect of execution than the top-level categorization (closed-loop, intermittent, open-loop), and rather than trying to empirically differentiate the effect of a single variable, it provides insight in the different tradeoffs that also involve other variables such as the inconsistency between display and motor space, the effects of warping objects in the visual field, and the addition of motor space.

The main questions motivating the experiment are:

- Which way of dealing with displayless space (ignoring it, or providing it without feedback) results in better CDOM performance? In which situations?
- What is the effect on performance of displayless space?
- Do off-screen feedback techniques help navigate displayless space?

6.3. Experiment 5: Mouse Ether vs. Stitching

Experiment 5 was designed to investigate the trade-offs involved in different ways of dealing with displayless in MDEs. Although the issue of execution is much broader that what can be explored with a single experiment, the results of this experiment are already directly applicable to current MDEs.

6.3.1. Apparatus

The experiment was conducted on a Windows™ Pentium IV™ machine connected to two identical 20” monitors (40.6cm by 30.5cm, 1600 by 1200px). The monitors were set at different heights (13.24cm offset) to control for the possible confound of corners (Mouse Ether could present an advantage due to the ability to *cut corners*, which is not possible with Stitching – see description of the techniques below for more detail). During the experiment, the monitors were moved between three different positions to test three different gaps between displays (Points A, B, and C in Figure 60). The monitors were arranged along a circumference of radius 90cm centered on the estimated position of the participant’s head; this kept the visual distance to the monitors constant for all gap distances and avoided possible perspective confounds (see [Nacenta et al., 2007b]).

The cursor was controlled by a Logitech G5™ Laser mouse, with a resolution of up to 2000 DPI. The experimental software avoided any quantization – a small movement of the mouse could produce a cursor movement of 1 pixel on the screen. Windows acceleration was disabled by using a flat curve in the corresponding Windows registry variable. The C-D gain was, therefore, constant at a value of 16 (640px/cm); a relatively high setting that avoided the need for clutching[18], even at the largest distances.

[18] In HCI terminology, *clutching* is the explicit disconnection of the tracking function of an input device. In a mouse, clutching involves raising the mouse from the surface to reposition it within the plane without registering any undesired movement. Clutching is what makes an indirect input device *relative.*

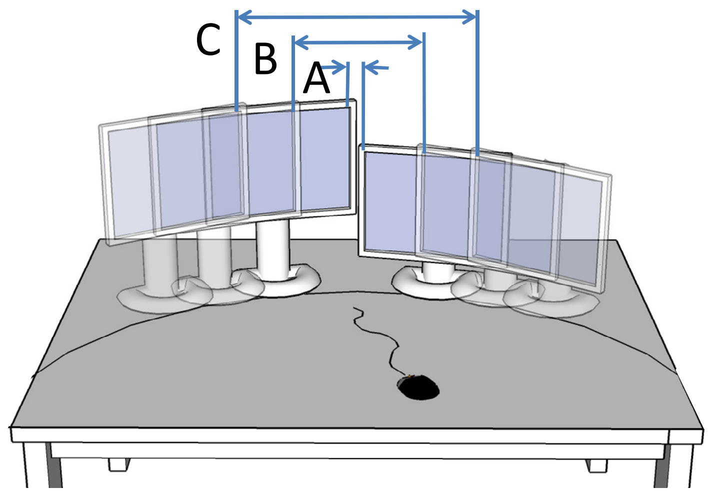

Figure 60. Experimental set-up with three different gaps between monitors.

Clutching was avoided because differences in clutching behavior across the experiment and between subjects could be a source of unwanted variability and produce confound effects (see [Casiez et al., 2008]). We disabled acceleration because it implies an extra source of motor-visual space mismatch that could negatively affect the performance of Mouse Ether.

The experimental software was built in C# for .NET 2.0. No delay was noticeable and the frame rate was above 24 frames/second.

6.3.2. Techniques

This experiment compared three techniques: Stitching (closed-loop), Mouse Ether (intermittent), and Mouse Ether with Halo (indirect closed-loop). A simulation of the techniques can be played through the accompanying video figure (Cross-display-cursor-movement.wmv).

6.3.2.1. Stitching

In Stitching the feedback loop is always available, but discontinuous in space due to the warp from one display to another. This implementation of Stitching replicates that of most current operating systems. The cursor jumps from one display to another when crossing a display boundary that is contiguous to some other display; if the cursor reaches a display boundary that is not connected to any other display, the cursor stays in the same display. When hitting the boundary, components of the input movement that are parallel to the screen boundary are still applied to the movement of the mouse (i.e., the cursor can still "slide" through the boundary – see Figure 61).

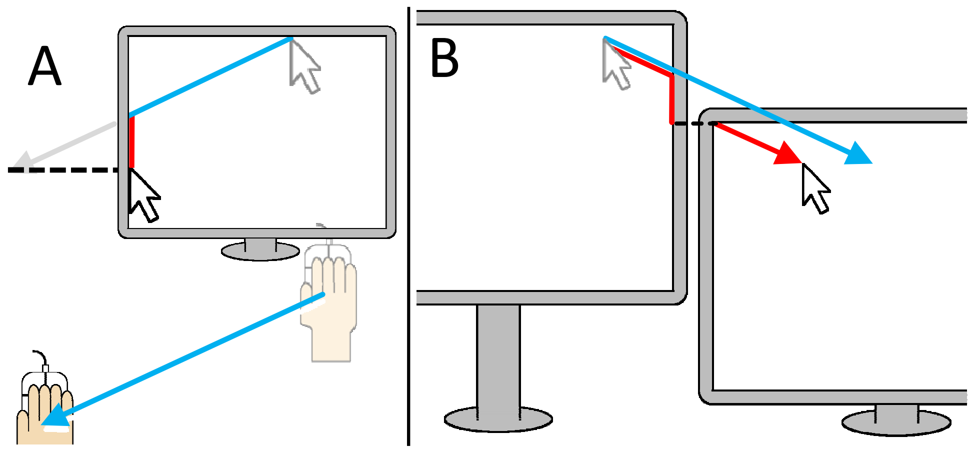

Figure 61. A) The outer boundaries of the display environment allow the cursor to "slide" (movement is right to left). B) The boundaries in Stitching prevent cutting corners and result in a partial loss of horizontal movement when sliding through the corner boundary (red arrow) with respect to the intended movement (blue arrow). Note that the vertical component of the movement is preserved because of the sliding mechanism.

In our implementation, the difference of alignment between screens is accounted for so that a perfectly horizontal movement of the mouse does not change the vertical position of the cursor when changing displays (Figure 62).

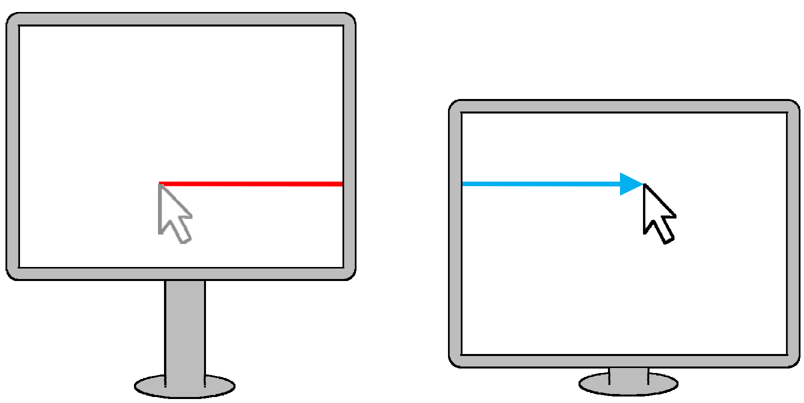

Figure 62. For all techniques, the vertical alignment was accounted for so that horizontal movements of the mouse produced horizontal cursor movements without height jumps.

As in the original implementation, the cursor could be moved in the ether until it reached the convex hull of the two displays (see Figure 63). Notice that this implementation restricts the cursor position but still allows cutting corners. Boundaries belonging to the convex hull behaved as in Stitching (i.e., the cursor would “slide” along the boundary as long as the mouse movement contained that directional component).

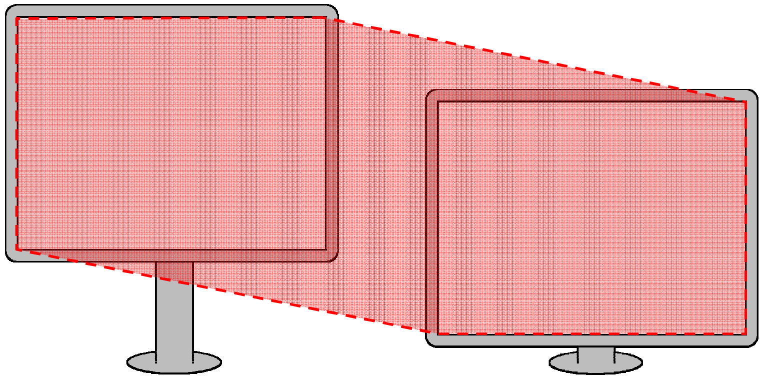

Figure 63. Limits of Mouse Ether as proposed by Baudisch and colleagues [5] (convex hull of display surfaces).

6.3.2.2. Ether+Halo

Ether+Halo behaved exactly like Mouse Ether except that a red circular halo appeared on screen when the cursor was in displayless space. The halo had a thickness of 3 pixels and was set to appear in at least one display at any time that the cursor was not visible. Halo intrusion borders (the area contiguous to the boundaries that can show a halo) covered a 200px (5cm) thick framework from the borders of each screen (see Figure 64).

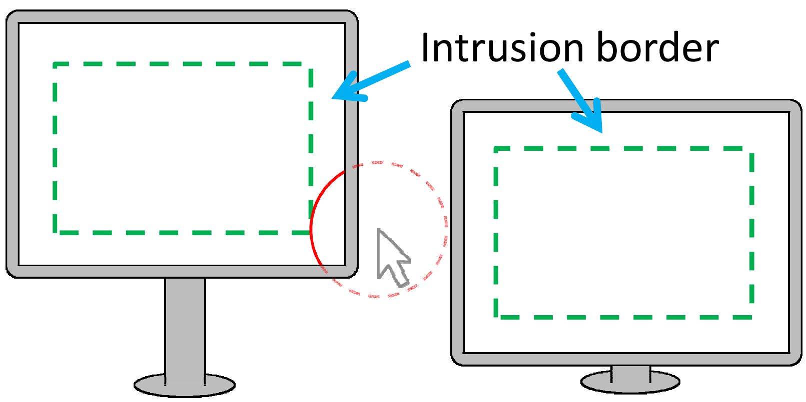

Figure 64. A halo (continuous red line) is always visible in at least one display when the cursor is in displayless space. The intrusion border is the display area that can show halos.

6.3.3. Participants

Twelve subjects – six females and six males, aged between 21 and 33 years – participated in the study for a $10 honorarium. Four participants had significant experience with multi-display environments and four had never used a multi-display system. The experiment took around 50 minutes to complete.

Participants signed a consent form and took a demographic questionnaire before the study, and filled a post-study questionnaire after. Samples of the consent form and the questionnaire are reproduced in Appendix A. Each subject was tested individually.

The experiment and the handling of participants were carried out according to the University of Saskatchewan's guidelines for experiments with human subjects.

6.3.4. Task

The task consisted of clicking in a blue square and then on a larger red square (4cm side, 158pixels). The initial and target squares were placed according to 10 possible paths designed to test different kinds of situations (see Figure 65).

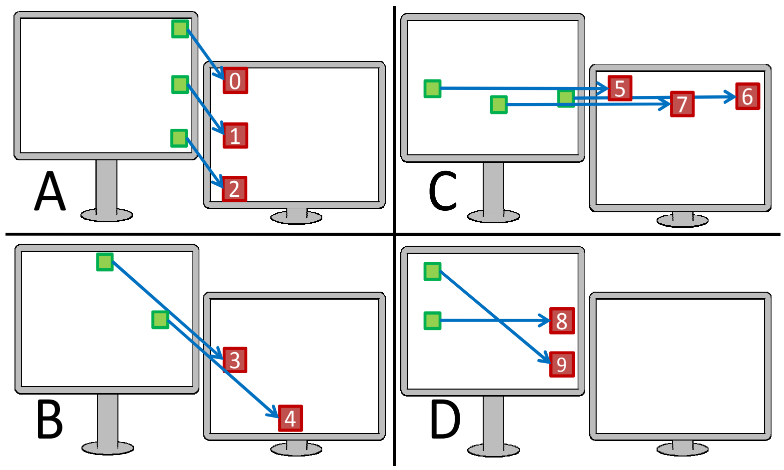

Figure 65. Paths used in the experiment. Note: representation is not to scale and Paths 4, 6 and 7 have a vertical offset to improve visibility in the figure.

All paths were left-to-right and diagonal paths were downward; pilot tests showed that completion times differ depending on direction, but direction did not interact with our variables of interest. Therefore it is reasonable to assume that right-to-left and bottom-top motions are affected similarly by cross-display technique and gap.

Paths 0 and 2 were designed to control for the effect of corners – rectilinear trajectories are possible with Mouse Ether, but not with Stitching. Path 1 is equivalent to 0 and 2 except that it does not come near any corner. Paths 8 and 9 represent single-display tasks and serve both as baseline comparisons between single- and multi-display performance and as distracters to avoid rhythmical completion of the tasks.

The main dependent variable of the study was completion time, which was measured from the initial click in the start square until the click in the target square. The software also registered errors, overshoots, times that the cursor entered the target, peak velocity, location of the peak velocity and time-stamped trajectories of the motion; these data, however, are only used to further understand relevant results from the planned tests.

6.3.5. Questions

The study was designed to answer four specific questions that were stated before the data analyses.

- Q1: Which cross-display technique is best for a particular distance between displays?
- Q2: Does Halo help performance in Mouse Ether?
- Q3: What is the relation between gap size and performance when using Stitching?
- Q4: Is the ability to cut corners an advantage of Mouse Ether over Stitching?

These questions correspond to specific statistical tests that were designed before data collection. The answers to these questions are linked to the specific research objectives of this chapter, formulated in Section 6.2.

6.3.6. Experimental Design

The experiment followed a 3x3x10 (three factor) within-subjects full-factorial design with cross-display technique, gap, and path as factors. The main factor was cross-display technique (Stitching, Mouse Ether, or Ether+Halo). The gap was either small (the minimum allowed by the monitors' bezels), medium, or large (positions A, B and C in Figure 8), which corresponded to gap angles of 2.4º, 14.4º and 26.4º (3.8cm, 23cm and 42cm along the circumference).

Participants were trained in each of the three cross-display techniques for each of the three possible gaps (30 trials per technique, 90 trials in total). They then performed blocks of 50 trials for each of the cross-display technique/gap combinations for a total of 150 trials per technique and a total of 450 trials in the main blocks.

The order of the inter-display mode condition was fully balanced across subjects (two participants in each possible order). Half of the participants performed the tests with increasing gaps (short, medium, long gaps) and half in the opposite order. In each block, the different paths appeared in a randomized order. The first 10 trials (one per path) of each block were marked as training and excluded from the main analysis to avoid noise and adaptation effects. Trials with errors (e.g., clicking outside the target) were repeated.

After all trials were completed, the participants filled in a post-study questionnaire with specific questions about preference, speed and accuracy of the cross-display techniques for each of the gap distances.

6.3.7. Results

The results are reported in three sections: the planned quantitative analyses that correspond to the main questions, the analysis of the subjective data from the post-study questionnaire, and the explanatory analyses.

All analyses were performed on error-free trials that were not marked as training. Four points per user for each cell (same cross-display technique, gap distance and path) were collected (a total of 3*3*10*4 = 360 valid data points per user).

6.3.7.1. Planned quantitative analysis

An omnibus factorial ANOVA with participant as random factor revealed that, as expected, there were significant main effects of cross-display technique ($F_{2,22} = 16.4$, $p < .001$), gap distance ($F_{2,22} = 283.1$, $p < .001$) and path ($F_{7,77} = 34.7$, $p < .001$). The interaction between cross-display technique and gap distance was also significant ($F_{4,44} = 13.6$, $p < .001$).

To answer question one (which cross-display technique is better at which gap distance?) and question two (does Halo help?) we performed planned pair-wise comparisons for each gap distance. All pair-wise comparisons in this section were corrected using the Games-Howell procedure for unequal variance data (analogous to Tukey's HSD).

Short Gap

Ether+Halo was the fastest technique (735ms average completion time), followed by Ether and Stitching (762ms and 768ms respectively). The pair-wise comparisons only show significant completion time differences between Ether+Halo and Stitching ($p < .008$) but this represents a small difference (33ms, 4.3%).

Medium Gap

Stitching was the fastest technique (865ms average completion time), followed by Ether+Halo (937ms) and Mouse Ether (995ms). All pair-wise comparisons were significant (all $p < .003$). The difference in completion time represents a performance advantage of Stitching of 7% and 13% with respect to Ether+Halo and Mouse Ether.

<u>Large Gap</u>

Stitching was again fastest (958ms), followed by Ether+Halo (1100ms) and Mouse Ether (1222ms). The ordering was the same as for the medium gap and the differences were also statistically significant (all p <.001), although proportionally larger; Stitching was 13% faster than Ether+Halo and 21% faster than Mouse Ether.

These analyses and Figure 66 allow us to give explicit answers to the first two questions:

Q1 (Which technique is better at which distance?): For the medium and large gaps, Stitching is superior to other techniques. For short gaps, the three techniques are roughly equivalent, although there is a slight advantage of Ether+Halo over Stitching.

Q2 (Does Halo help?): Ether+Halo was faster than Ether at all distances. The effect of Halo is not statistically significant for the short gap, but the improvements of Halo on movement time are statistically significant for the other two gap distances.

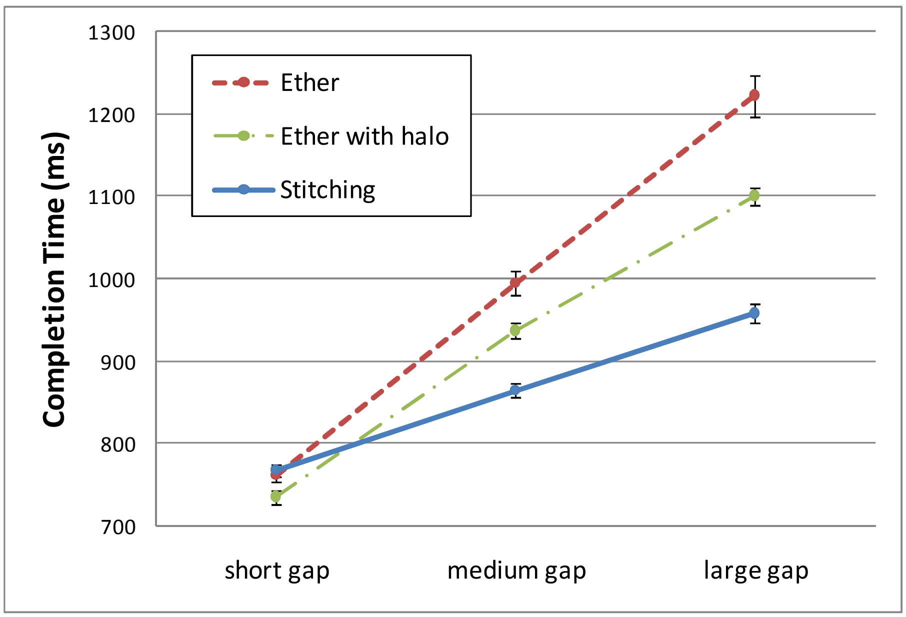

Figure 66. Completion time by gap and technique, Paths 8 & 9 excluded (error bars display standard error). Note that the y axis starts at 700ms.

Table 9. Average completion time (in milliseconds) of the different techniques (rows) according to the size of the gap between displays (columns). Numbers between parentheses represent standard error.

	Short Gap	Medium Gap	Large Gap
Stitching	768 (7)	865 (9)	958 (12)
Ether	762 (8)	995 (14)	1222 (25)
Ether+Halo	735 (8)	937 (10)	1100 (11)

To answer question three – relationship between gap distance and performance for Stitching – two comparisons were planned; short to medium gap and medium to large. Both comparisons showed significant differences ($p < .001$). This relationship is easiest to see in Figure 67, which compares targeting times in Stitching for Paths 3 and 4 with Path 8, which is 4 cm longer but does not cross the gap.

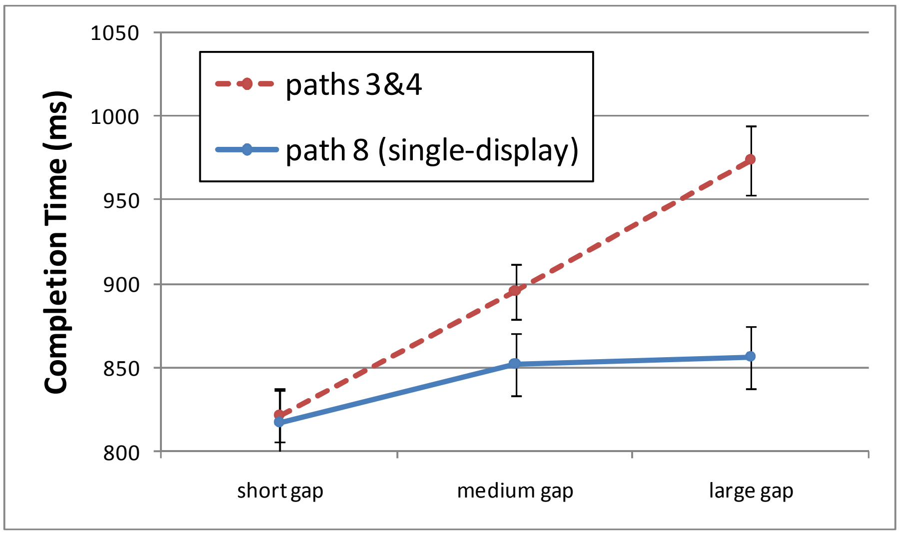

Figure 67. Completion time by gap for Paths 8, 3 and 4 with Stitching (error bars display standard error). Note that the y axis starts at 800ms.

Table 10. Average completion time (in milliseconds) of two sets of tasks of the Stitching technique with a similar targeting length (3&4, and 8 - rows) according to the size of the gap between displays (columns). Numbers between parentheses represent standard error.

	Short Gap	Medium Gap	Large Gap
Paths 3 and 4 (cross-display)	821 (15)	895 (16)	973 (20)
Path 8 (single-display)	817 (20)	851 (18)	856 (18)

Q3 (What is the relationship between gap and completion time with Stitching?): Although the distance traveled by the mouse in motor space is constant for the Stitching technique at all gaps, the completion time increases with the size of the gap.

Question four is answered by looking at a comparison between Paths 1 and 3 (which might benefit from cutting corners) and Path 2 (which does not require cutting any corner). In our analysis, all cross-display techniques show similar completion times for Paths 1, 3 and 2. A factorial ANOVA with participant as a random factor and using data exclusively from Paths 1, 2 and 3 did not show any interaction between path and cross-display technique ($F_{2,22}$ = 3.423, p >.05).

Q4 (Is cutting corners an advantage of Mouse Ether?): no significant improvement in targeting times was observed between techniques in the corner paths, contradicting the hypothesis that cutting corners provides a significant advantage for Mouse Ether. Moreover, these paths show the same pattern as the rest: Stitching is faster or indistinguishable from other techniques.

Across the experiment, errors (i.e., clicking outside targets) were below 6% for any given technique-gap combination; errors were evenly distributed across techniques (5.3% Stitching, 4.7% Mouse Ether, 5.5% Mouse Ether+Halo).

6.3.7.2. Subjective Data

The post-study questionnaire asked participants to rank the three cross-display techniques for each of the gaps according to three criteria: personal preference, speed, and accuracy (a total of 9 rank sets). Non-parametric Friedman analyses for each of the gap-criteria combinations show a marginally significant difference only in the short gap ($\chi^2(2)$ = 6.16, p <.046) out of the nine tests. No α-level adjustments were performed.

Responses from participants varied widely, as is reflected in their comments; some participants found that Mouse Ether was "more natural" and that "the cursor comes up where I expect it to be". However, other participants stated that "[Stitching] is faster" and that "[in Mouse Ether] without the halo I became lost once or twice". Some participants liked the halo because "[it] helped me get unlost" while others found it "annoying", "distracting" and "confusing", especially for the short gap.

Remarkably, some participants stated that "[with Stitching] I felt lost sometimes" and "the distance is still too far and seems to jump too fast".

6.3.7.3. Explanatory Analyses

During the study we logged several other measures besides completion time that can help explain our findings by revealing user behavior in the context of models for targeting. Using the dependent measures of peak velocity, time to peak velocity, percent time after peak velocity, display where peak velocity occurs, and presence of an overshoot error, we conducted factorial ANOVAs on each dependent measure that offered insights on the results for Q3 and Q2; we report these in this order for clarity.

Q3 (What is the relationship between gap and completion time with Stitching?): Our main findings show that as the physical distance between displays increases, there is a corresponding increase in completion time, even though the physical movement of the mouse remains constant. Fitts's Law [MacKenzie, 1991; Graham and MacKenzie, 1995] predicts movement time based on a user's movements in motor space, not on the distance traveled by the cursor. Although we can assume that switching attention between displays during targeting has a time cost, is this reflected in a slower ballistic movement or a lengthening of the homing-in phase? Previous studies have shown that the timing and magnitude of the peak velocity in a targeting task are a function of movement amplitude regardless of precision constraints, and that the proportion of time spent decelerating to the target is a function of an accuracy constraint, (usually seen through small target widths), regardless of the distance covered [Graham and MacKenzie, 1995].

Our analyses on the data from Stitching revealed that participants reached higher top speeds ($F_{2,22} = 5.02$, $p < .05$) with increasing gap and spent longer in the ballistic phase ($F_{2,22} = 33.1$, $p < .001$). However, the higher speed did not help overall performance. In addition, with larger gaps, participants were not spending a greater proportion of time in the homing-in phase ($F_{2,22} = 1.41$, $p > .05$). Thus, longer movement times are not a result of a slower ballistic phase or a longer deceleration phase; participants were actually moving faster in conditions with larger movement times.

An increase in top speed and time to reach top speed are indicative of longer movement amplitudes, revealing that participants planned and executed the ballistic phase of the movement based on the physical configuration of the displays, even though they were aware that the cursor would immediately warp to the second monitor. In fact, users reached their top speed on the destination display 39% of the time for a small gap, but 58% and 57% of the time for the medium and large gaps respectively. As a result, we can expect that users will consistently overshoot the target when aiming over a gap, which explains the longer movement times. An examination of the percentage of trials with a horizontal overshoot of the target showed that overshoot errors increased dramatically with a physical gap. For the small gap, 36% of trials contained an overshoot as compared to 68% for the medium gap, and 80% for the large gap. These additional overshoots came as a result of users incorporating the physical space between monitors into their motor planning, which subsequently increased movement times when warping the cursor over the gap.

Q2 (Does Halo help?): Our main findings show that Halo reduced movement time when aiming across displayless space. In our study, users were only without visual feedback within the gap, creating a closed-loop to open-loop to closed-loop transition within a single aiming movement. The two-phase model of movement suggests that if users are in the ballistic phase of their movement within displayless space, the halo will not be helpful, and might slow a user down by forcing them into closed-loop aiming. If users are decelerating within the gap, the halo should help because it provides essential visual feedback for movement correction.

Examining only the Mouse Ether and Mouse Ether+Halo data, we used peak velocity to determine on which display the ballistic phase ended and the feedback-corrected homing-in phase began. For trials where the ballistic phase ended prior to the gap, or within the gap, there was a movement time advantage with Halo for the medium and large gaps, but not for the small gap. For trials where the ballistic phase ended on the destination display there was no advantage of Halo for the small or medium gap. For the large gap, there was a difference; however there were only 48 trials (less than 2% of analyzed trials) that comprised this sample for comparison.
In addition, there were no effects of Halo on peak velocity ($F_{1,11}$=2.14, p>0.05), time to reach peak velocity ($F_{1,11}$ = 2.47, p >.05), or percent time spent in the deceleration phase ($F_{1,11}$ = 1.01, p >.05). Halo did result in decreased time after peak velocity ($F_{1,11}$ = 7.71, p <.05), likely due to a decreased number of overshoots. Examining the horizontal overshoots showed that there was a smaller percentage of trials with an overshoot when Halo was used for the medium gap (Ether: 36% Halo: 30%), and the large gap (E: 53% H: 35%), but not the small gap (E: 25% H: 23%). These results suggest that the halo is indeed used by participants to know where the cursor is in displayless space, assisting with targeting by decreasing overshoots, thus decreasing movement time.

6.3.8. Discussion

6.3.8.1. Stitching vs. Mouse Ether

The results of this experiment place Stitching as the technique with the best performance at all distance gaps except at the shortest distance. The similarity of the techniques' performance at the shortest distance is hardly surprising, since when the gap is small, the three techniques are almost equivalent. However, the superiority of Stitching over Mouse Ether overall is by no means trivial; in fact these results contradict those from the initial evaluation of Mouse Ether [Baudisch et al., 2004].

The contradiction can be explained by the differences in implementation of the Stitching technique in Baudisch and colleagues' experiment: their version of Stitching did not compensate for the differences in resolution between the two displays, and this factor might create the disadvantage that the more general experiment reported here did not show (see also explanation in Section 6.1.4.3, page 113, and Figure 59).

It is likely that the advantage shown by Stitching is due to the shorter motor distance. However, there seems to be considerably more going on in terms of feedback and spatial consistency that can be accounted for only by the targeting distance. For example, the improvement in performance when adding the halo suggests that the lack of feedback during a section of the targeting has a negative effect on targeting time.

In terms of the goals of this chapter, these results indicate that Stitching (closed-loop, inconsistent motor-visual space – warping –, and shorter movement amplitude) is preferable to Mouse Ether techniques (consistent motor-visual space, indirect or no feedback in displayless space, longer movement amplitude), at least in the simple (but common) two-monitor scenario that was tested. In other words, even though trying to preserve continuous feedback (closed-loop) produces discontinuity in its spatial representation, the lack of feedback in the middle of the targeting motion (intermittent), added to the extra motor space required to make the visual and motor space consistent, are more detrimental that the warp.

6.3.8.2. The Impact of Visual Space vs. Motor Space

The second most relevant result is that the warp of closed-loop techniques causes a loss in performance that increases as the displayless space grows between displays (about an extra 7ms per angular degree of displayless space in between displays). This is true even when the motor distance between targets is kept the same. The result is relevant because it shows that displayless space should be minimized when possible, and that when considering targeting performance it is important to look at how the visual feedback is presented, and not only at how the task maps into motor space.

The explanatory analysis of the data of Section 6.3.7.3 hints at an explanation that would summarize the results from this and the previous subsection: ignoring the displayless space negatively affects the ballistic part of the targeting motion because, even after training, participants are not able to dissociate the motor planning from the visual information of the physical configuration. At the same time, the lack of feedback in a consistent motor-visual space (intermittent) negatively affects the homing-in phase of the movement, enough to make the technique the slowest of the pack. From this data alone we cannot extrapolate that slowing down the ballistic phase (mostly open-loop) is generally better than affecting the homing-in phase (heavily reliant on feedback, and therefore mostly closed-loop), but this explanation seems plausible.

It is worth highlighting that even though the actual target distance for Stitching remains the same across all gaps (Figure 66), and is always shorter than for the ether techniques, Mouse Ether is faster at the short gap than Stitching at the medium gap, even though the targeting distance was larger. Similarly, the medium-gap-halo is approximately equal to the large-gap-stitch, even though the targeting distance was considerably larger. These findings attest to the importance of considering displayless space when designing CDOM techniques.

6.3.8.3. Off-screen Feedback

The study included Mouse Ether+Halo, a condition in which spatial consistency between the motor and visual space was kept, and feedback was still present through the whole movement (if through an indirect medium – the halo). The data reveal that, although giving indirect feedback helped, it did not help enough to make Mouse Ether+Halo the most efficient technique. This might be due to the discontinuity in the representation of the feedback – switching between different ways of signaling the position of the cursor has an associated performance cost, at least with the Halo off-screen technique.

The participant feedback also indicated that, for some people, the halo might be distracting or obtrusive. This problem might be exacerbated in situations with several users or when there is real data on the screen (not only the targeting elements as in this experiment). This does not mean that off-screen feedback should be discarded; although Halo was the technique to beat at the time of designing this experiment, future and existing redesigns of off-screen feedback (e.g., [Gustafson et al., 2008]) might result in improved performance.

6.3.8.4. Further Improving Performance

It is unclear if new auxiliary techniques can be devised that could reduce the loss of performance due to the gap in Stitching. Possible candidate strategies include slowing down the cursor at the entrance of the destination screen to avoid overshooting (an approach similar to [Mandryk et al., 2005; Mandryk and Gutwin, 2008]), and providing visual aids to guide gaze to the cursor more effectively after the warp.

6.3.8.5. Cutting Corners

As a minor result, the results reveal that boundaries (corners) do not have a significant negative effect. We know boundaries can be useful for other things. For example, boundaries facilitate some targeting tasks because elements on the boundary become infinitely deep in motor space [Accot and Zhai, 2002]. Current interfaces take advantage of this property and place frequently-used elements on the boundaries (e.g., the task bar in Windows XP and Vista™ or the menu bar in Mac OS X™). Mouse Ether eliminates some of these hard boundaries in exchange for the ability to cut corners; the analysis of Experiment 5 could not, however, find any significant advantage of cutting corners. It is probably better not to sacrifice these boundaries for extra 'ether' around screens if they can be used to facilitate targeting of frequently used objects.

6.3.8.6. Limitations of the Study

The results of this study cannot be generalized yet to all kinds of MDEs. The scenario was limited to displays that were aligned at the same distance of the participants and perpendicular to the line of gaze. Although this scenario is probably the most common, there is some evidence that the warping of the cursor might affect performance and/or user preference when the screens are at a sharper angle from each other (see Chapter 5, and [Su and Bailey, 2005]). The results from Experiment 4 (Section 5.3.5.1) show that Stitching across similarly sized displays is slower than Perspective Cursor, a technique that could be considered a generalization of Mouse Ether+Halo for complex 3D environments.

Also, mouse acceleration[19] was not included in any of the conditions of this study. Although we can only speculate on the effect of acceleration, there is some evidence that the effect of acceleration is generally small [Casiez et al., 2008], and, even if it were important, it would probably only further increase the performance difference (mouse acceleration alters the consistency between the motor and the visual space, which happens to be the most important conceptual advantage of Mouse Ether).

[19] Mouse acceleration is a non-linear dynamic manipulation of the Control to Display ratio that is applied in some current operating systems. The basic concept is that the cursor will cover a longer screen trajectory when the mouse is moved fast than when it is moved slowly over the same physical distance.

6.4. Discussion

This section looks at the empirical findings of Experiment 5 in the broader context of the goals of Chapter 6 and the previous literature.

At the beginning of this chapter I divided interaction techniques into open-loop, closed-loop and intermittent according to the presence of a feedback loop. Although this chapter did not include open-loop techniques in any empirical comparison, we know from the literature on motor control and related experiments in HCI that open-loop techniques trade speed for accuracy, resulting in fast but imprecise results. It can be informative to further study comparisons of open-loop approaches with closed-loop and intermittent techniques. It could be useful as well to develop open-loop techniques that provide reasonable accuracy at no additional cost in speed, although these techniques will be probably useful for a limited set of scenarios.

In any case, the study of open-loop techniques is secondary to this research because open-loop techniques are not likely to be affected by the lack of feedback specific to MDEs (open-loop techniques are open-loop because they lack this feedback), and because most of current multi-display interaction still requires precision (e.g., the placing of windows across different displays, the dropping of an icon into a particular interface widget and the selection of a tool from a toolbar located in another monitor).

The results of this experiment do, however, shed light on how the discontinuity of the feedback loop, the spatial discontinuity of feedback and the size of displayless space affect performance when targeting across displays. The results and analysis of existing techniques show a complex interdependence of these factors. Partial lack of feedback and inconsistencies between motor and visual space both negatively affect performance, but when the choice is available, it is probably best to avoid the lack of feedback, even if this means warping the object from one screen to the next.

The source of the advantage of Stitching over Mouse Ether cannot be completely attributed to the continuous presence of feedback because to make the motor and visual space consistent with each other in Mouse Ether it is necessary to increase the motor distance; however, as discussed in Section 6.3.8.1, the improvement shown on Mouse Ether when the Halo was used implies that lack of feedback is a problem and has, at least, a partial effect. This issue deserves further study, but it is important to remark that there are currently no alternatives to adding motor space in order to achieve visual-motor consistency. Therefore, regardless of the relative importance of each of the two possible sources of performance loss (extra motor space or lack of feedback), the two cannot be separated, and the alternatives that MDE designers will face are the same that were studied in this experiment[20].

One way to avoid the decision between Stitching-like and Ether-like techniques altogether is to select literal techniques, which, by their very nature, are always closed-loop. This option is, however, accompanied by the tradeoff between power and efficiency that was amply discussed in Chapters 4 and 5.

It should also be noted that it is still early to fully generalize the findings of this chapter to any MDE, task, or input type. Some configurations might exacerbate the problems seen with the warping of the cursor (or objects) over the limitations of the lack of feedback (as seen, e.g., in Mouse Ether), or the problems of switching between different forms of feedback (e.g., from direct feedback to off-screen feedback). Although further research will be required to corroborate the relationships between the three input configurations for other tasks and other types of interaction techniques, the empirical results presented here form a consistent basis that identify the problems with cross-display movement execution, present the alternatives for design, and provide evidence supporting the use of closed-loop techniques, at least for the most common scenarios.

6.5. Implications

There are five main lessons from this chapter that can translate into advice for the design of CDOM interfaces:

- If high accuracy is not required, open-loop techniques (e.g., flick) can provide faster interaction
- If accuracy in CDOM is required, open-loop is probably not an alternative, and closed-loop or intermittent techniques will likely perform better
- Performance in CDOM tasks with closed-loop techniques (e.g., Stitching, Mouse Ether+Halo) and intermittent techniques (e.g., Mouse Ether) is negatively affected by

[20] Notice that, even though it might be possible to manipulate the C/D ratio to make the distance between targets in motor space the same for Mouse Ether and Stitching, this manipulation also alters the relationship between the distance and the target, and, according to Fitts's Law, will result in the same performance. The lack of an effect of C-D gain changes on performance has been confirmed in [Casiez et al., 2008] within a certain range of values.

the discontinuity of feedback and/or the extra motor space required to make the motor and visual space consistent

- Some fully-closed-loop techniques (such as Stitching) provide the advantage of reduced targeting distance (through warping), although this advantage should be measured against the performance degradation due to the inconsistency between motor and visual space.
- When choosing between not presenting feedback between the displays (intermittent, e.g., Mouse Ether), presenting varying sorts of feedback (closed-loop, e.g., Mouse Ether+Halo), and presenting continuous feedback that warps in space (closed-loop, e.g., Stitching), the latter provides better performance, at least in MDEs with relatively simple spatial configurations

6.6. Conclusions

In this chapter I have identified the control paradigm of the interaction technique as an important factor that affects cross-display object movement performance in multi-display environments. According to their control paradigm I classify existing techniques into open-loop, closed-loop and intermittent.

Because existing research shows that open-loop control is fast but inaccurate, and because the lack of feedback of these techniques is not specific to MDEs, I decided to focus on the performance differences as affected by different ways of dealing with displayless space. A comparison between two closed-loop techniques (Stitching and Mouse Ether+Halo) and an intermittent technique (Mouse Ether) showed an advantage for closed-loop techniques; however, the design issues involved go beyond the presence or lack of feedback and involve also the discontinuity of visual-motor feedback and the addition of extra motor space to achieve the latter. Although the results are not definitive enough to separate the performance according only to the control paradigm, the findings offer important information for designers on how to deal with the fractured nature of MDEs. In practical design terms, these results suggest that the best way to deal with displayless space is to simply ignore it, even if this implies that the cursor warps between one display and the next.

Another important result is that either way of dealing with displayless space (ignoring it, or accounting for it with or without the lack of feedback) negatively affects performance. This effect increases with the size of the gap between displays.

CHAPTER 7: DISCUSSION

This chapter analyzes and discusses the combined findings from the studies presented in the previous four chapters. A summary of the findings is provided first, followed by discussion and explanation of the results. I then discuss the scope and limitations of what has been learned and of the methods used. The chapter finishes with a summary of lessons for practitioners and a discussion and critical reflection on the relevance, and real world impact of this work.

7.1. Summary of Findings

In the previous four chapters I have described five experiments that investigate the effects of different types of techniques on the cross-display object movement performance. Experiments 1 and 2 look at the planning part of the CDOM action; in particular, at how two different ways of referring to the destination (spatial and non-spatial) affect human performance. Experiments 1 and 2 show that laying out the interface in a way that corresponds to the physical world can facilitate the selection of a display; specifically, Experiment 1 shows that this is the case when the task is itself spatial (i.e., when the stimuli comes in a spatial form), whereas Experiment 2 shows that performance can also be affected if the task is not directly connected to the physical position of the stimuli – that is, the physical layout of objects (e.g., displays and people) has an effect on performance even if the task is not spatial, at least when learning associations between real and virtual elements is required.

Experiments 3 and 4 further explore spatial techniques; in particular, they explore how the relationship between the control action required by a technique and the physical layout of the space can affect performance and preference. CDOM techniques are classified into three groups (planar, perspective, and literal), which correspond to different relationships between the physical configuration of displays and the input model that the interaction techniques use. The two experiments uncover differences in both performance and preference between the three technique groups. Experiment 3 compares a range of planar techniques (techniques in which the input model represents the physical layout of displays as a flat surface) and literal techniques (techniques where the input and the location of the display are in the same place – e.g. Pick-and-

Drop) and shows that literal techniques are more efficient for CDOM; however, literal techniques are always limited by reach, whereas planar and perspective techniques are not. Experiment 4 compares planar and perspective techniques (which are not bound by reach limitations), and shows that matching the control actions of the user with their perception of the physical space (the main characteristic of perspective techniques) improves performance for MDEs that have displays of different sizes in different orientations.

Experiment 5 focuses on the execution of CDOM actions. In particular, the experiment analyzes the different ways in which displayless space can be dealt with. The issue of displayless space is specific to MDEs, and affects several other important factors (e.g., presence of a feedback loop, motor distance, motor-visual consistency) that are also important for performance. The findings show that ignoring displayless space, even at the cost of spatial consistency between motor and visual space, results in better performance than providing motor-visual consistency through increasing the motor space (making feedback discontinuous, and increasing the targeting distance), at least for simple monitor settings.

7.2. Overall Discussion

The results from Experiments 1 and 2 show that the relationship between the spatial arrangement of displays in physical space and the control actions required to select a display can affect performance of both spatial and non-spatial tasks. The results from Experiment 1, in which a spatial task was tested, are not surprising; it seems evident that aligning the interface with the layout of the physical world will probably result in faster actions, since the users' cognitive processes are simplified (the mapping between stimuli and response are straightforward). However, this has never been measured in the context of HCI tasks; generally, the advantage of spatial mappings in the interface is assumed (e.g., [Hazas et al., 2005]) without empirical evidence of the existence or size of the advantage. This is surprising because creating interfaces that are aware of the physical space requires considerable sensing, effort, engineering, and expense.

The results from Experiment 1 provide support for the design of interfaces that are *compatible* with the physical reality of what they control: the advantage exists and Experiment 1 calibrated it to a magnitude of close to 30%, which, as discussed in Section 3.3.7, is probably a lower bound of the effects that we can expect when the experimental constraints are removed and when more users and displays are involved. The results of this experiment can help designers make the right decision about what kind of CDOM techniques to support; however, the ultimate decision of whether the effect is large enough to warrant spatial tracking of the physical positions of people and objects rests on the designer of the system, who has to evaluate whether the difference in performance is worth the cost and complexity of the extra sensing equipment. These decisions are only possible with a deep understanding of the larger task and the frequency with which CDOM actions will take place in the specific scenario of use of the system.

The results of Experiment 2 are much less evident than those for Experiment 1, but also less strong. Findings from Experiment 2 suggest that spatiality is an integral part of co-located interaction, even if the task itself is not spatial. Although the differences found were not as large as for Experiment 1 and seem to be important only during the learning phase of the task, these findings can have important implications for the design of co-located systems, since they imply that, regardless of the relationship of the task with the physical locations of objects, spatially-aware interfaces can show better performance. As with Experiment 1, it is likely that the differences found between spatial and non-spatial layouts of the interface are only a lower bound on the actual differences that we would observe in more realistic settings, where compensating the natural impulse to react spatially is harder because CDOM tasks are not likely to happen in consecutive order, with clear feedback, and without interruption, as they did in the experiment.

The results from both Experiment 1 and 2 point to the construction and design of interfaces that are more aware of the spatial location of actors and objects in the space. Although as humans we are capable of great abstraction, we are also spatial beings, and the location and layout of the environment around us should be taken into account for the design of new technology. This is supported by current research in Ubiquitous Computing (e.g., [Abowd and Mynatt, 2000]), which tends to expand the sensing around users to make systems that are capable of reacting without the need for explicit user interaction. However, the results from Experiments 1 and 2 suggest that this sensing is also needed to make explicit interaction more efficient and fluid for the users.

What was found is also largely consistent with phenomena and theories from psychological research. In Chapters 3, 4 and 5, results are linked to previous research on Stimulus-Response Compatibility and the more generic Dimensional Overlap model. Although I believe that it is important to study new specific scenarios that relate to the use of computers, it is also crucial for the field of HCI to relate these findings back to more general phenomena, models and theories studied in other fields such as Psychology and Cognitive Science. This is important not only to avoid reinventing the wheel, but also to enable generalization of results to new situations and to contribute to a more accurate understanding of how design must be influenced by mental, environmental and human constraints.

Experiments 3 and 4 take the investigation of spatial techniques further by looking at the specific relationships between the actual physical layout of the displays and the input that is required by the interaction techniques to produce cross-display object movement. The analysis of interaction techniques that led to the experimental design of Experiment 3 contributes as much information as the results from the experiment itself. Experiment 3 shows how literal techniques (those that require direct contact with the objects that are being moved or manipulated) are more efficient, but they restrict the power of users (i.e., how far and with which accuracy they can reach a certain object in a distant part of a display). Restricted power is often not an option, and we also know through other research that users will prefer not to move away from their current locations to interact with objects (e.g., [Voida et al., 2005]); therefore it becomes important to find out which techniques will allow CDOM at a distance in an efficient way. The results from Experiment 3 are useful for determining which design alternatives among the wide range of possible planar interaction techniques can provide power and performance simultaneously. One of the options is to provide a miniature representation of the display layout so that users can conveniently manipulate objects in the within reach space (e.g., Radar). This is an attractive solution and was shown by Experiment 3 to be the most efficient; however, as was discussed in Sections 5.2.9.2 and 5.4, performance might be negatively affected by switching attention from the miniature to the real objects and back. Miniatures are also problematic if the representation of the object makes it difficult to identify the specific object that needs to be moved, for example, because the resolution of the miniature is not enough to discriminate the object of interest from other objects.

If world-in-miniature techniques are not feasible for a particular environment, it might still be possible to use other planar indirect interaction techniques such as the Pantograph, Slingshot, or Telepointers (as in the applications described in [Nacenta et al., 2007]). The results from Experiment 3 indicate that the performance of planar techniques is inferior to literal techniques;

moreover, it seems likely that with a more complex layout of displays, these techniques might result in interfaces that are unnecessary difficult to operate.

These observations prompted the study of a different class of techniques (perspective techniques) and the design and execution of a new experiment (Experiment 4) that would look at a more complex multi-display environment.

Perspective techniques are techniques in which the location of the user with respect to the environment is used to improve the mapping between how the environment is perceived and how the input is provided (see Section 4.1.2). Perspective techniques are not completely new; the most common of these, Laser Pointing (a.k.a. Distant Pointing or Laser Beam) has often been considered for MDEs and large displays [Olsen and Nielsen, 2001; Oh and Stuerzlinger, 2002; Parker et al., 2005; Nacenta et al., 2007]. However, it has always been difficult to classify Laser Pointing with other techniques: sometimes it was considered a direct technique, sometimes an indirect technique. The perspective category captures the dynamic nature of Laser Pointing (the mapping depends on the location of the device) but it also allows the classification of techniques from other areas that would otherwise have to be ignored (e.g., techniques based on eye or head tracking [Benko and Feiner, 2005; Ashdown et al., 2005] and augmented reality [Butz et al., 1999; Hill and Johnson, 2008]). The concept of perspective techniques also enabled the design of a novel CDOM interaction technique that is based on mouse control but uses the position of the eyes to establish the relationship between mouse movement and physical location of the cursor.

The results of Experiment 4 confirm that performance can be improved without compromising power by using perspective techniques that take into account the user's point of view for the mapping of input. The results show solid evidence for the case of relatively complex MDEs, such as the one designed for the experiment; however, indications from other experiments not described in this dissertation suggest that the perspective approach can also generally improve performance for cursor movement within large and oblique displays [Nacenta et al., 2007b], and not only complex MDEs.

It is also important to note that, within the perspective category, there is still ample room for variation. Different techniques that create different mappings, and the use of different input devices can also significantly affect performance and preference. Experiment 4 showed relevant differences between Laser Pointing and Perspective Cursor that can be due to the input device (the mouse has long been proved to be a very efficient device [Card et al., 1978], and laser pointers are notoriously difficult to stabilize [Myers et al., 2002]), or to the fact that the input-perceptual match is better in Perspective Cursor, where the “perspective” corresponds to what the user sees, and not to what the pointer “sees”, as with the Laser Pointer technique. The accuracy of the match between input and point of view of the user also brings us back to the Dimensional Overlap model, which predicts that a better match between perception and action will result in easier, faster and less error-prone interaction [Kornblum, 1992].

In more general terms, the results from Experiment 4 confirm the support that Experiments 1 and 2 offer for the design of interfaces that are spatially aware of the environment that we want to control; however, these conclusions can be taken even further through the findings of Experiment 4, which suggest that systems can be more efficient not only if they know about the physical environment of the user, but also if the interface adapts depending on how the environment is perceived by the user. Fortunately, this extra step of “making the interface put itself in the shoes of the user” is a matter of just a few extra geometrical calculations once the appropriate sensing hardware is available. Other experiments (not described in this dissertation) indicate that the benefits can extend beyond CDOM into other areas such as recognition and perception of shapes and reading [Nacenta et al., 2007b].

Experiment 5 closes this series of studies of CDOM by looking at the third element in the CDOM process, the level of execution. MDEs are fragmentary by nature because visual output is constrained to specific surfaces (displays), and it is generally not possible to display information in the space between displays (displayless space). This fragmentation of output creates a dilemma that concerns the crucial information loop between the interface and the user: we can either present spatially continuous feedback that is interrupted during the transition between one display and the next, and that requires the mouse to traverse this extra motor space, or we can provide temporally continuous feedback that warps from one display to another. The experiment tested several of the alternatives: traditional Stitching ignores displayless space but provides visual feedback at all times, Mouse Ether provides spatial consistency, but with temporal interruptions of feedback and added motor space, and the Mouse Ether+Halo tries to overcome the temporal lack of feedback with the Halo off-screen feedback technique.

The results from Experiment 5 indicate that, for simple MDEs (the kind that are already present in a large proportion of PC settings, especially in offices), interrupting feedback in the middle of CDOM and adding extra motor space to create consistency between the spatial and visual spaces results in poorer performance than simply ignoring displayless space, even if it is at the cost of warping the cursor (i.e., of producing discrepancy between input and visual spaces). These results seem to go against the general trend of the four previous experiments, which indicate that taking into account the point of view of the user pays off in terms of performance and accuracy. A more detailed look at the results shows, however, that not preserving spatial consistency (as the ignored amount of display space grows) has a negative effect in performance (in line with the general explanations of the other experiments), but the combined negative effect of the added motor space and the lack of feedback (or of presenting different kinds of feedback at different points during the movement process) is much stronger than the benefit of spatial consistency, at least when the MDE is very simple as in Experiment 5. It is possible that in more complex environments (such as that used in Experiment 4) the effects of neglecting spatial consistency would overpower the problems that come from forcing this consistency; namely, the extra motor space and the temporal discontinuity of the feedback. This is supported by some of the results in Experiment 4, which show that Perspective Cursor (which has to "travel" across the displayless space between the display surfaces of the setting of Figure 47 – in page 83), was still significantly faster than Stitching, even for display transitions that only differ in the angle at which displays are oriented with respect to each other.

The current data from Experiment 5 cannot completely explain the source of the disadvantage of the intermittent and open-loop techniques. We know that a larger targeting distance should result in longer targeting times (according to Fitts's Law), but the experiment also showed that providing off-screen feedback (in this case, a halo) significantly speeds up the task, even after accounting for the likely time cost of switching back between sources of feedback (direct visual feedback vs. Halo).

There are two lessons that we can derive from the combined analysis of Experiments 4 and 5. First, suspending the feedback loop in a CDOM action, introducing inconsistency between motor and visual space, and adding motor space to a task have significant effects in the performance of CDOM tasks. Second, the relative importance of these three factors very likely depends on the complexity of the MDE and the geometrical relationships between displays. The results of Experiment 5, however, also have implications that go beyond the particular technique that is used for cross-monitor cursor transitions. In particular, these findings remind us that the predictions of Fitts's Law for targeting are also affected by what the user sees, not only by how far or with how much precision she has to move her limbs.

Some researchers have suggested that the problem of how to deal with displayless space might not be relevant in the near future, since displays and projectors are becoming larger and cheaper and might soon cover all space around us without fractures, or will be worn in glasses, allowing digital content to overlap any space in our environment. It is difficult to know which direction the design of new interfaces is going to take, but I believe it is dangerous to predict a future human-computer interface of not only ubiquitous, but also fully immersive displays around us. CAVEs [Cruz-Neira et al., 1993] have been very popular in research environments since the 1980s but they have failed to deliver benefits as work and entertainment environments; there are very few people willing to spend prolonged periods of time interacting in claustrophobic spaces covered by displays. Instead, it seems that the emerging MDEs are increasingly a combination of many displays of different sizes and characteristics, such as modern meeting rooms where a vertical projector has to share attention with the displays of laptops, netbooks, PDAs, smartphones and digital media players.

Together, the five experiments provide a cross-cutting analysis of CDOM that goes from the more generic distinctions between spatial and non-spatial techniques and how people can best refer to displays, to the more concrete specifics of the execution of an interaction technique. Although the empirical exploration presented in this dissertation cannot be comprehensive, it presents evidence on several design decisions that, according to the results from the experiments, can substantially affect performance and preference.

These design decisions also generate categories into which we can classify existing CDOM techniques. In particular, the first level divides techniques into spatial and non-spatial according to how displays are referred to – that is, according to their referential domain (Chapter 3) – the second level divides spatial techniques into literal, planar and perspective according to their input model (Chapters 4 and 5), and the third level divides techniques into closed-loop, open-loop and intermittent according to the presence of feedback (Chapter 6). This classification can be considered a first taxonomy of CDOM interaction techniques (depicted in Figure 68).

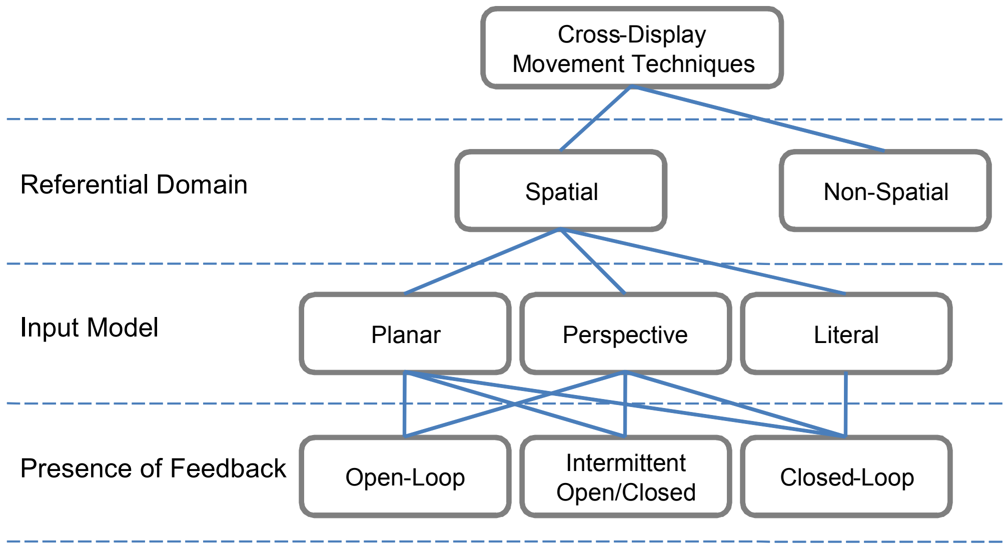

Figure 68. A taxonomy of CDOM techniques.

This taxonomy focuses on characteristics that affect performance and preference. It is possible and necessary to extend it with new categories under the existing ones (e.g., sub-categories of perspective techniques), and even to create orthogonal levels that reflect other design decisions that affect other important aspects of interaction techniques in MDEs, such as awareness, conflict, interference and monetary cost. Nevertheless, the taxonomy of Figure 68 (expanded and detailed in Figure 69 and Appendix B) can already be useful for several reasons: it classifies existing CDOM techniques into categories that can help explain differences in preference and performance; it summarizes interdependencies between the different levels (e.g., literal techniques can only be closed-loop); and it exposes areas that are relatively unpopulated and in which exploration could result in new types of techniques (e.g., alternative mappings under the non-spatial category).

7.3. Scope and Limitations

The research described in this dissertation is necessarily limited by constraints in breadth, depth, methods, and time. In this research (which is itself supported by careful work by others in the areas of HCI and Psychology) I have tried to build a solid base that could help designers to understand the state of the art in CDOM, and assist other researchers in bringing our understanding of the issues of MDEs further. This section extends Section 2.1 (page 8) to make explicit the limits of the analysis and results described in the previous chapters, and to indicate the areas where research is most urgently needed to answer important questions. The obvious consequence that follows from most of the remarks in this chapter is future research, which will be discussed in Section 8.4 (page 156) of the concluding chapter.

The following subsections discuss three main areas of the limitations and scope of this work: the coverage of the design space of CDOM techniques, the limitations in measures and methodology employed, and the questions that cannot be answered through the experiments.

7.3.1. Coverage of the Design Space

The three main design decisions explored in this dissertation are part of a much larger space and there are other possible parameters in interaction techniques that might yield interesting designs. Some of the explorations and discussions from previous chapters provide hints as to what these could be; for example, world-in-miniature techniques have only been explored partially (Chapter 4), and it might be possible to cross the WIM approach with some of the other elements of the design space presented here or in external work. For example, it might be advantageous to create versions of Radar that have different mappings to reality, even adapted to the user's perspective.

Non-spatial interaction techniques are another potentially large unexplored area. Although there has been some work that suggests mappings of the interface other than the representation of spatial location, such as mappings directly created by users [Aliakseyeu and Martens, 2006; Aliakseyeu et al., 2008], there are still numerous unexplored mapping possibilities such as relating the location and size of elements to their type, the frequency of use, etc. The possibilities are almost infinite, but should probably be explored in the context of specific applications, since the effectiveness of a particular mapping (not as generic as spatial location) will depend on whether it is useful for the task at hand.

Another area of possible exploration and sub-categorization are perspective techniques. There are many unexplored ways of introducing the user perspective as a dynamic part of the input mapping, and some of them might yield better results for separate tasks.

7.3.2. Measures and Methodology

Constraining the evaluation of techniques mostly to performance and user preference allows for a very specific focus and the possibility of direct comparison of techniques. Performance is an important measure because it is highly desirable that the atomic task of CDOM can be finished fast and without error. The underlying assumption is that CDOM in itself is usually not the primary goal of the user's activity with the system; therefore, the faster that CDOM and other atomic operations can be completed (e.g., opening a file [Collins et al., 2007]), and the less effort they take, the less it will interrupt the natural flow of the higher-level task.

However, interaction techniques can also affect other important variables that form part of the operation and goals of MDEs. For example, it has been shown that territorial behavior (the spatial distribution of activity by the different users around the available space of the environment) can be affected by the design of an interaction technique; more specifically, by the visual design of the technique [Nacenta et al., 2007; Pinelle et al., 2008]. Alternative measures of interest that can be found in the groupware literature include awareness [Gutwin and Greenberg, 2002], equality of participation [Rogers et al., 2009; Marshall et al., 2008], interference and conflict [Hornecker et al., 2008], and engagement and fun [Mandryk, 2004]. To these we can add some obvious ones that have not yet been formalized in research such as implementation complexity and monetary cost.

The measures chosen for this research also influenced the choice of methods employed. The main research method, the controlled experiment, is a good match for performance because it can be measured quantitatively and with less noise. Controlled experiments also have the advantages of reliability and replicability, and provide specific answers for the hypotheses and questions posed. However, controlled experiments also have several limitations: they are necessarily narrow and specific in the questions that can be asked; require a careful control of the environment and the design of the task, which can constraint the external validity and generalizability of the task if not done carefully; and they are often not suited for certain types of research [Greenberg and Buxton, 2008]. As the study of CDOM interaction techniques progresses, it will necessarily have to expand the applied methodologies; specifically, as we need to learn how often and why CDOM takes place and the effects of certain techniques on collaboration, researchers will have to move away from controlled quantitative methods into less controlled and more qualitative observations (for a review on the range of methods available for HCI and Computer Supported Collaborative Work see [McGrath, 1995]).

7.3.3. Unanswered Questions

As I pointed out across the previous chapters, there are questions and issues that, even within the limited scope of this dissertation, are still left without answers because experiments had to focus on specific topics or because these questions were generated from the results of this research. This subsection compiles the open questions that have been mentioned in the discussion sections of Chapters 3, 4, 5 and 6, and contributes some new ones that arise from an overall look at the research.

7.3.3.1. More Destinations and Different Focuses

One of the unanswered questions that arise from the two studies in Chapter 3 is whether the presence of more displays or possible destinations will affect spatial and non-spatial techniques differently. Variations in the number of displays or people in Experiments 1 and 2 were not part of the experimental design. As mentioned in Sections 3.3.7, 3.4.7, 3.5, and 7.2 there are reasons to believe that effects will only be accentuated with more people and more displays, but we do not yet know to what extent. Similarly, the controlled nature of Experiments 1 and 2 does not give any indication of how these effects will play out in an environment in which the main goal is not only to perform cross-display object movements, but rather concerns higher level issues such as finishing a document, carrying out a successful meeting, or making an argument.

7.3.3.2. The Planar-Perspective Comparison

Chapters 4 and 5 present the categorization at the level of input model into three categories: literal, planar and perspective. Experiment 3 compares several planar and literal techniques while Experiment 4 compares a planar technique with two perspective techniques. Although the two experiments offer a number of interesting comparisons (also with respect to the different input alternatives of CDOM techniques) the performance comparison between perspective techniques and literal techniques is missing. Although extrapolating from the experiments of this dissertation and external research suggests that literal techniques are difficult to beat whenever targets are within hand's reach, having empirical data on this comparison would probably be useful for designers. It would also be valuable to determine the critical point at which users would chose to use a powerful but less efficient technique over one that requires more physical effort.

7.3.3.3. Interactions Across Design Decisions

The standard scientific approach to complex problems such as CDOM requires that we first divide the larger issue into smaller sub-problems that can be attacked separately. Here the secret of good and useful research is a careful and intelligent division of the problem so that separate results still hold when several variables across sub-problems change simultaneously.

Unfortunately, there can be no certainty that the problem division chosen is not subject to interactions between variables across different sub-problems, neither is there a method that can be applied to get optimal results.

In the case of this research, the division has been made according to an analysis of the CDOM process, but this provides no guarantee of lack of interaction; it is possible that certain variables of design, especially at the level of input model and presence of feedback, interact with each other in unpredictable ways. This does not detract from the results achieved; instead it suggests that, as with any other scientific inquiry, it requires further validation (and possibly, reconsideration) of the results.

7.3.3.4. Worlds-in-miniature

The discussions of Chapter 4 and the results and limitations of Experiment 3 open interesting questions as to the use and value of world-in-miniature techniques for CDOM. WIM techniques provide an alternative visual representation of elements in the physical or the digital environment that can be very flexible and designed to facilitate interaction; for example, making the visual representation smaller (a miniature, as is suggested in its name) makes it easy to have an overview of the whole space of interest at a glance and oriented to the advantage of the user, and allows interaction everywhere, since the whole interactive representation falls within arm's reach. In exchange, the duplication of the space carries problems of its own: in order to manipulate the "real" object through the miniature we have to first find the object in the miniature that represents it; the miniature representation might not be sufficient in itself for manipulation (e.g., pictures where visual details are important); and the duplication of space implies that the space is no longer shared with other users when interaction happens mostly through the miniature [Nacenta et al., 2007].

When are miniatures easier to manipulate than the actual objects? What are the most important parameters of the design of miniatures? What are the collaborative effects of the availability of replicated spaces? These are all important questions that do not have yet definitive answers, but are beginning to be considered in the HCI literature (e.g., [Gutwin, 1997; Pinelle et al., 2006; Nacenta et al., 2007]). Although CDOM is only one of the many aspects relevant to WIM techniques, CDOM techniques can greatly benefit from an improved knowledge of the human factor issues of worlds-in-miniature.

7.3.3.5. The Role of Mouse Ether

The results of Experiments 4 and 5 showed some contradictions with respect to Mouse Ether and with respect to the effect of introducing extra motor space to increase the consistency between the motor and the visual spaces. In Experiment 4, the Perspective Cursor condition (which includes a perspective version of Halo) showed better performance than Stitching (which simply ignores displayless space) in the *simple transition* tasks (between a horizontal display and a vertical display of approximately the same size and separated by a gap – see Figure 47, in page 83, and Figure 52.A, in page 90). Incidental observations of the participants executing the task indicate that with Stitching people often overshot the target, whereas they did not with Perspective Cursor. In Experiment 5, overshoots were also common for the Stitching condition; however, Stitching did outperform Mouse Ether+Halo. In both experiments, the addition of "Ether" meant the increase of motor space between the targets.

There are several possible explanations for this discrepancy. Since the main difference between the Ether+Halo of Experiment 5 and the Perspective Cursor of Experiment 4 is the perspective (angular) mapping, it is possible that the advantage comes only from the perspective mapping, even for simple transitions. Alternatively, we might be observing one of the interactions described in Section 7.3.3.3 (page 147) and Ether might be useful for transitions across displays that are at an angle. It is not possible to determine from the data of these experiments which one is the source of this discrepancy because the experiments were not designed to answer this question.

7.3.3.6. Hybrid vs. Pure Techniques and the Redundancy of Interaction Techniques

Experimental design is limited by constraints such as limited participant time. A common mistake in naïve experimental designs is to include too many factors or too many conditions, which weakens the power of the experiment and makes it very likely to fail to see differences, even if the effects are large. This limitation is often avoided by a careful choice of conditions (or techniques) that are representative of pure design alternatives. However, this usually leaves out hybrid designs which combine characteristics from several "pure" techniques.

It is impossible to empirically investigate all hybrids because the number of combinations quickly increases with just a few alternatives of a few design decisions. I believe that the solution to this conundrum lies partly in the practice of interface design (designers can analyze and consider hybridizations of techniques if they have a good understanding of the basic characteristics of “pure” techniques), but it could also be helpful to investigate more generically the conditions in which hybrid techniques are likely to provide extra value, and in which they just become a liability.

This discussion is parallel to that of the inclusion of redundant techniques in a single interface. It is often the case that different techniques are best for certain subtasks; whether several redundant techniques are included should be weighed against the increase in complexity and the added selection time that the multiplicity of techniques can cause. More research in this area is needed, but I believe that it is probably better to apply an Occam’s Razor argument to the design of interfaces: do not provide redundant techniques if tasks can be reasonably accomplished with a small set.

7.3.3.7. The Relationship between Performance and Preference

In all the experiments except for Experiment 5, the subjective evaluations of techniques followed the objective measures of performance. In Experiment 5 the subjective evaluations were inconclusive as to whether participants had preference for any of the techniques. The correlation between performance and preference could well be an artifact of the controlled laboratory study approach, but it also suggests that, at least at the level of execution of small atomic actions, performance is preferred over any other characteristic of the technique.

Parallel to the discussion in previous subsections, the experiments carried out for this dissertation provide only weak evidence to suggest that preference follows performance; however, this is actually a relevant question to answer in future research and has deep implications for the design of interfaces: what are the other factors or technique characteristics that make a technique preferable over more efficient techniques? To what extent are subjective evaluations of techniques (generally easier and cheaper than empirical measures) accurate in predicting the performance of a technique?

7.4. Lessons for Practitioners

Chapters 3, 5 and 6 already contain specific design guidelines interpreted from the results of each of the five experiments. This section summarizes some of the most important guidelines suggested by the results of the experiments.

- Spatial CDOM techniques are preferable to non-spatial techniques, at least when the extra hardware that might be required is not an obstacle.

- Spatial interaction techniques should be considered even for tasks that are not necessarily spatial.
- Literal techniques can provide the fastest interaction, but are limited in power.
- World-in-miniature techniques can be efficient and powerful, but these advantages should be weighed against the possibility of added transition time between the miniature and the actual manipulated objects.
- In complex MDEs, techniques that take perspective into account can provide better performance than more naïve planar models and with more power than literal techniques; however, the benefits should also be measured against the extra cost and complexity of making a system perspective-aware.
- The space between displays causes delays in tasks that take place across displays; if possible, space in between displays should be minimized.
- For systems with relatively simple geometry, it is preferable to provide constant visual feedback than to provide spatial consistency of input and feedback at the cost of adding motor space.

7.5. Relevance, Potential Real-world Impact and Critical Reflection

Multi-display environments are already common. Many desktop PCs are connected to several monitors, and laptops are often used with extra monitors. MDEs are likely to soon be even more common and include more displays because adding displays seems to provide evident advantages with almost no drawbacks: it provides extra screen real estate (which reduces scrolling and switching between windows), and different displays connected to the same machine can be used for different purposes (e.g., for sharing information with audiences in public presentations, or to split the output application according to the current focus of work [Grudin, 2001]).

Besides multi-monitor systems, MDEs are also widespread in specialized contexts such as control rooms, audio-video production facilities, and smart offices in research labs [Johanson et al., 2001; Ni et al., 2006; Nunamaker et al., 1991; Rekimoto and Saitoh, 1999; Tandler et al., 2001; Wigdor et al, 2006; Wigdor et al, 2009]. Although these will likely become gradually even more common, making their efficient design a relevant matter, there is also a large unexplored potential in turning existing multi-computer environments in which displays are loosely connected (e.g., offices with many PCs, homes with multiple appliances – for games, music, information, domotic control), into MDE interfaces that work in an integrated fashion. One of the most obvious steps to improve the interoperability of displays is to provide easy cross-display object movement. Object transfer across groups of computers is notoriously cumbersome; co-workers usually resort to USB keys or e-mail in order to transfer files and some people even prefer to print a document or copy it by hand just to avoid the hassle of electronic transfer [Rekimoto]. Interestingly, this is not a only problem of lack of infrastructure or networking technology: fast networks, wired and wireless, are ubiquitous and already allow for fast file transfer. One of the basic elements missing, however, are interfaces that are useful from the point of view of the user.

This dissertation work is a step in that direction. By examining the CDOM process in detail and analyzing several factors that can play an important role on performance and preference, this work expands our understanding of interaction techniques for CDOM. The analysis and data presented in Chapters 3 to 6 can help designers create interfaces that enable easier and more efficient transfer of objects between displays, which in turn, can improve the efficiency of interaction across displays, but also facilitate collaborative activities that are currently hindered by cumbersome interfaces. This work can also inform the HCI research community, call attention to issues that have not been studied in depth, and establish a common vocabulary to describe phenomena and factors that are important in this domain.

Naturally, the work presented in this dissertation is necessarily limited in scope and by the inherent constraints of available technology. It is also important to note that, although the findings are meant to be useful for designers who need to introduce CDOM techniques in their MDE systems, this work does not intend to be a comprehensive examination of all the possible design alternatives of CDOM interaction techniques.

It is difficult to track and foresee how and to what extent the findings here described will influence the future of MDE interfaces; nevertheless, the potential is large, given the large number of existing MDEs and the probable growth of these systems in the near future. Moreover, many of the results that have already been published in the form of several conference papers and a journal article have been picked up by others, and some work is already being developed by other researchers and by me that builds upon this basis. It is my hope that these results help designers and researchers better understand human perception, cognition and motor behavior as it applies to the design of CDOM interaction techniques and that, eventually, this work can help the collaborative effort of improving the interfaces of the work tools of the future.

CHAPTER 8: CONCLUSION

Cross-display object movement is a fundamental action of multi-display environments. The ability to transfer objects from one display to another enables flexible interaction with all the available display space of an MDE system and allows users to carry out their activities with the display that is most adequate for the particular task. Little multi-display interaction is possible without cross-display object movement. In this dissertation I focused on the study of CDOM interaction techniques, which are the actual interactive mechanisms that enable CDOM.

8.1. Research Objectives

In this research I set out to improve our knowledge of factors that affect human performance on cross-display object movement tasks. In particular, I tested the hypothesis that *human performance in cross-display object movement actions is affected by the way that the cross-display interaction technique refers to destination displays, by the way that the interaction technique maps the physical environment, and by the way that displayless space is dealt with during the execution of the movement.*

In the preceding five chapters I have presented and discussed evidence that shows that each of the three factors implicit in the hypothesis (the way of referencing the space, the mapping of the physical environment, and the way to deal with displayless space) do produce significant differences in the performance of cross-display object movement tasks.

These research results provide much-needed evidence on how different design decisions can result in interaction techniques with substantial differences in performance. Designers of future multi-display environment interfaces and interaction researchers can apply these results to choose techniques that are appropriate to the purpose of their systems, or to design new interaction techniques that are based on the improved understanding of how cross-display object movement works.

8.2. Main Contributions

These are three main contributions provided in this dissertation:

- Through an analysis of the cross-display object movement process I have identified three factors that affect performance in CDOM actions: how displays are referred to, how the physical environment is mapped in the interface, and how displayless space is deal with during the execution of the action. These factors have not been considered before in HCI or Psychological literature.
- Through a series of experiments, I have provided new empirical evidence that shows that the three factors identified result in performance differences, and an estimation of the magnitude of these differences.
- The analysis of the factors generate different design options that result in different types of interaction techniques. Different ways of referring to displays result in spatial and non-spatial techniques. Different ways to map the physical space result in planar, perspective and literal techniques. In the case of execution, we considered techniques with closed, intermittent or open feedback loops, but looked further than this categorization into different ways to deal with displayless space: introducing it in the motor space (Ether techniques), ignoring it (Stitching techniques), or providing indirect visual feedback for it (with the help of off-screen feedback techniques).

8.3. Minor Contributions

As minor contributions I provide:

- A categorization of existing techniques into a taxonomy derived from the design options discussed above.
- A set of recommendations for practitioners on how to choose techniques and what choices to avoid in multi-display environments.
- Perspective Cursor: a new perspective mouse-based technique for multi-display interaction.
- Perspective Halo: an adaptation of Halo for off-screen location of objects in non-planar environments.
- Mouse Ether+Halo: an adaptation of Halo for the purpose of tracking a cursor when traveling in displayless space.
- A model of the cross-display object movement process.

8.4. Future Research

Future research in the area must start by addressing the open questions discussed in Section 7.3:

- How does the number of displays affect CDOM performance?
- How do planar techniques compare to perspective techniques?
- Do unexpected interactions exist between the three studied factors?
- What are the characteristics of WIM techniques for CDOM in real scenarios?
- When is Mouse Ether useful?
- Which hybrid interaction techniques can be beneficial, and when is it useful to provide redundant techniques?
- What is the relationship between performance and user preference in CDOM interaction techniques?

Since this research focused primarily on performance measures (a fundamental first step to understand how to improve CDOM), it is also important to extend the research to include other measures and factors that can also be relevant. For example, it is relevant to learn about how different techniques can affect group work processes (an area that we have already started to explore [Nacenta et al., 2007; Pinelle et al., 2008; Pinelle et al., 2009]).

The analyses, categorizations, and empirical tests described in the previous seven chapters have also uncovered a number of possibilities for the design of new techniques. Future research can take advantage of this work by exploring techniques from relatively unexplored areas such as new variants of non-spatial techniques, world-in-miniature adaptations, and perspective techniques.

Finally, some of the topics studied for this dissertation may lend themselves to further refinement in the form of quantitative models. For example, the effect of displayless space discussed in Chapter 6 could be modeled within a Fitts's Law-style equation. Similarly, it might be possible to predict targeting times across displays parameterized by angle, distance and size of the origin and target displays. Quantitative models can be very useful because they compile knowledge in a concise form that can be used for design without the need of new empirical testing; however, the data gathered in this work is not yet sufficient to formulate models.

8.5. Conclusion

Multi-display environments are increasingly common and will likely become even more popular in the future. To realize the potential benefits of MDEs (e.g., large amounts of screen real-estate, environments suited for multi-user collaboration, and flexible use of displays that match specific tasks), we need to be able to present interfaces that are useful and efficient, and that necessarily depart from current interface design approaches that were designed with the monitor-keyboard-mouse setup in mind.

In this dissertation I have analyzed and explored one of the actions that are specific to MDEs and that is fundamental to support flexible multi-display interaction: cross-display object movement. This dissertation provides evidence on how decisions in the design of interaction techniques can affect performance in cross-display object movement, categorizes these decisions in groups that lead to a taxonomy of interaction techniques, and describes a novel interaction technique for multi-display environments that is based on the idea of perspective.

This effort builds upon previous research and provides useful information for future research towards the ultimate goal of producing more efficient and useful computer tools for human use.

REFERENCES

Abowd, G. D. and Mynatt, E. D. 2000. Charting past, present, and future research in ubiquitous computing. *ACM Trans. Comput.-Hum. Interact.* 7, 1 (Mar. 2000), 29-58.

Accot, J. & Zhai, S. More than dotting the i's - foundations for crossing-based interfaces. 2002. In *Proceedings of the SIGCHI Conference on Human Factors in Computing systems, CHI'02*, ACM, New York, NY, 73-80.

Agarawala, A. and Balakrishnan, R. 2006. Keepin' it real: pushing the desktop metaphor with physics, piles and the pen. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada, April 22 - 27, 2006). R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, and G. Olson, Eds. CHI '06. ACM, New York, NY, 1283-1292.

Aliakseyeu, D. & Martens, J.-B. (2006) Sketch Radar: A Novel Technique for Multi-Device Interaction. *Proceedings of HCI'2006, Vol. 2*, 45-49. British HCI Group

Aliakseyeu, D., Lucero, A., Martens, J.-B. (2008). Users' quest for an optimized representation of a multi-device space. In *Personal and Ubiquitous Computing*, vol. 13, n. 8. Springer London, 599-607.

Asano, T., Sharlin, E., Kitamura, Y., Takashima, K., and Kishino, F. 2005. Predictive interaction using the delphian desktop. In *Proceedings of the 18th Annual ACM Symposium on User interface Software and Technology* (Seattle, WA, USA, October 23 - 26, 2005). UIST '05. ACM, New York, NY, 133-141.

Ashdown, M., Oka, K., and Sato, Y. 2005. Combining head tracking and mouse input for a GUI on multiple monitors. *Extended Abstracts of the CHI'05 Conference on Human Factors in Computing Systems,* 1188-1191. New York: ACM

Apple Corporation. *Setting up multiple displays as an extended desktop.* http://docs.info.apple.com/article.html?path=Mac/10.4/en/mh798.html Last accessed June 7th 2008.

Ayatsuka, Y., Matsushita, N., and Rekimoto, J. 2000. HyperPalette: a hybrid computing environment for small computing devices. *Extended Abstracts of CHI '00 Conference on Human Factors in Computing Systems*, 133-134. New York: ACM

Balakrishnan, R. 2004. "Beating" Fitts' law: Virtual enhancements for pointing facilitation. *International Journal of Human-Computer Studies, 61(6),* 857-874.

Baudisch, P., Cutrell, E., Hinckley, K. and Gruen, R. 2004. Mouse Ether: Accelerating the Acquisition of Targets Across Multi-Monitor Displays. *Extended Abstracts of the CHI'04 Conference on Human Factors in Computing Systems,* 1379-1382. New York: ACM.

Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B., and Zierlinger, A. 2003. Drag-and-pop and Drag-and-pick: Techniques for accessing remote screen content on touch- and pen operated systems. *Proceedings of INTERACT 2003 - the Ninth IFIP TC13 International Conference on Human-Computer Interaction*, 57-64. Amsterdam: IOS Press.

Baudisch, P. and Rosenholtz, R. 2003. Halo: a technique for visualizing off-screen objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA, April 05 - 10, 2003). CHI '03. ACM, New York, NY, 481-488.

Benko, H. and Feiner, S. 2005. Multi-monitor mouse. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (Portland, OR, USA, April 02 - 07, 2005). CHI '05. ACM, New York, NY, 1208-1211.

Benko, H. and Feiner, S. 2007. Pointer warping in heterogeneous multi-monitor environments. In *Proceedings of Graphics interface 2007* (Montreal, Canada, May 28 - 30, 2007). GI '07, vol. 234. ACM, New York, NY, 111-117.

Bezerianos, A. & Balakrishnan, R. 2005. The Vacuum: Facilitating the manipulation of distant objects. *Proceedings of the CHI'05 Conference on Human Factors in Computing Systems,* 361-370. New York: ACM.

Biehl, J. T. and Bailey, B. P. 2004. ARIS: an interface for application relocation in an interactive space. In *Proceedings of Graphics interface 2004* (London, Ontario, Canada, May 17 - 19, 2004). ACM International Conference Proceeding Series, vol. 62. Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo, Waterloo, Ontario, 107-116.

Biehl, J. T.; Bailey, B P. 2006. Improving Interfaces for Managing Applications in Multipe-Device Environments. *Proceedings of AVI '06 Conference on Advanced Visual interfaces*, 35-42. New York: ACM.

Biehl, J. T., Baker, W. T., Bailey, B. P., Tan, D. S., Inkpen, K. M., and Czerwinski, M. 2008. Impromptu: a new interaction framework for supporting collaboration in multiple display environments and its field evaluation for co-located software development. In *Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy, April 05 - 10, 2008). CHI '08. ACM, New York, NY, 939-948.

Bolt, Richard A. 1980. "Put-that-there": Voice and gesture at the graphics interface. *Proceedings of SIGGRAPH '80 Conference on Computer graphics and interactive techniques*, 262-270. New York, ACM.

Booth, K. S., Fisher, B. D., Lin, C. J., and Argue, R. 2002. The "mighty mouse" multi-screen collaboration tool. In *Proceedings of the 15th Annual ACM Symposium on User interface*

Software and Technology (Paris, France, October 27 - 30, 2002). UIST '02. ACM, New York, NY, 209-212.

Britton, E. G., Lipscomb, J. S., and Pique, M. E. 1978. Making nested rotations convenient for the user. *SIGGRAPH Comput. Graph*. 12, 3 (Aug. 1978), 222-227

Butz, A., Höllerer, T., Feiner, S., MacIntyre, B., and Beshers, C. 1999. Enveloping Users and Computers in a Collaborative 3D Augmented Reality. *Proceedings of the 2nd IEEE and ACM international Workshop on Augmented Reality IWAR' 1999*, 35-44. Washington: IEEE Computer Society.

Butz, A. & Krüger, A. 2006. Applying the Peephole Metaphor in a Mixed-Reality Room. *IEEE Computer Graphics and Applications*, *Jan/Feb, 2006, vol. 26, no. 1,* 56-63

Buxton, W. 1983. Lexical and pragmatic considerations of input structures. *SIGGRAPH Comput. Graph.* 17, 1 (Jan. 1983), 31-37.

Card, S. K., English, W. K., and Burr, B. J. Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. Ergonomics 21, 8 (Aug. 1978), 601-613.

Card, S. K., Mackinlay, J. D., and Robertson, G. G. 1991. A morphological analysis of the design space of input devices. *ACM Trans. Inf. Syst.* 9, 2 (Apr. 1991), 99-122.

Card, S. K., Mackinlay, J. D., and Shneiderman, B. Eds. 1999 *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers Inc.

Card, S. K., Newell, A., and Moran, T. P. 1983. *The Psychology of Human-Computer Interaction*. L. Erlbaum Associates Inc.

Carlton, L.G. Visual processing time and the control of movement. In *Vision and Motor Control*. North Holland, Amsterdam, 1992, 3-31.

Casiez, G., Vogel, D., and Balakrishnan, R. 2008. The impact of control-display gain on user performance in pointing tasks. *Human-Computer Interaction* 23, 3 (2008), 215-250.

Chiu, P., Liu, Q., Boreczky, J., Foote, J., Tohru, F., Kimber, D., Lertsithichai, S., Liao, C. 2003. Manipulating and annotating slides in a multi-display environment. In *Proceedings of INTERACT'03.* 583-590.

Chua, R., Weeks, D. J., and Goodman, D. 2003. Perceptual-motor interaction: some implications for human-computer interaction. In J. A. Jacko & A. Sears, (Eds.), *the Human-Computer interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications. Human Factors And Ergonomics* (pp. 23-34). Mahwah, NJ: Lawrence Erlbaum Associates.

Cockburn, A. & Firth, A. 2003. Improving the acquisition of small targets. In *Proc. British HCI Conference 2003*. 181-196.

Collins, A., Apted, T., and Kay, J. 2007. Tabletop File System Access: Associative and Hierarchical Approaches. Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems, TABLETOP '07., 113-120.

Collomb, M., Hascoët, M., Baudisch, P., and Lee, B. 2005. Improving drag-and-drop on wall-size displays. *Proceedings of the GI'05 conference on Graphics interface*, 25-32. New York: ACM

Crabtree, A., Hemmings, T. and Rodden, T. 2004. The social construction of displays. In O'Hara, K., Perry, M., and Churchill, E. (eds.) *Public and Situated Displays: Social and Interactional Aspects of Shared Display Technologies (Cooperative Work, 2)*. Kluwer Academic Publishers. pp. 170-190.

Crossman, E.. & Goodeve, P. 1983. Feedback control of hand movements and Fitts' law. In *Quarterly Journal of Experimental Pschychology*, 35, (1983), 251-278. Taylor and Francis.

Cruz-Neira, C., Sandin, D. J., and DeFanti, T. A. 1993. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In *Proceedings of the 20th Annual Conference on Computer Graphics and interactive Techniques* (Anaheim, CA, August 02 - 06, 1993). SIGGRAPH '93. ACM, New York, NY, 135-142.

Davis, J., and Chen, X. 2002. LumiPoint: Multi-user laser-based interaction on large tiled displays. *Displays 2002,* 23(5):205-211.

Dickie, C., Hart, J., Vertegaal, R. and Eiser, A. 2006. LookPoint: an evaluation of eye input for hands-free switching of input devices between multiple computers. *Proceedings of OZCHI06, the CHISIG Annual Conference on Human-Computer Interaction*. pp. 119-126.

Dulberg, M.S., Amant, R.s. and Zettlemoyer, L.S. 2003. An Imprecise Mouse Gesture for the Fast Activation of Controls. *Proceedings of INTERACT 2003 - the Ninth IFIP TC13 International Conference on Human-Computer Interaction*, 57-64. Amsterdam: IOS Press

Elliott, D., Binsted, G., and Heath, M. 1999. The control of goal-directed limb movements: Correcting errors in the trajectory. In *Human Movement Science* 18, 2-3 (1999), 121-136.

Fitts, P.M. 1992 (reprinted - 1954). The information capacity of the human motor system in controlling the amplitude of movement. In *Journal of experimental psychology. General* vol.121 iss.3 p.262.

Fitts, P.M. and Deininger, R. L. 1954. S-R compatibility: Correspondence among paired elements within stimulus and response codes. *Journal Experimental Psychology, 48,* 483-492.

Fitts, P.M. & Peterson, J.R. 1964. Information capacity of discrete motor responses. *Journal of Experimental Psych*ology, 67, (1964), 103-112.

Fitts, P.M. and Seeger C.M. 1953. S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal Experimental Psychology, 46*, 199-210.

Fjeld, M., Voorhorst, F., Bichsel, M., Lauche, K., Rauterberg, M., Krueger, H.. 1999. Exploring Brick-Based Navigation and Composition in an Augmented Reality. In *Handheld and Ubiquitous Computing*. *LNCS 1707*. 102-116.

Forlines, C., Esenther, A., Shen, C., Wigdor, D., and Ryall, K. 2006. Multi-user, multi-display interaction with a single-user, single-display geospatial application. *Proceedings of UIST 2006 – the ACM Symposium on User Interface Software and Technology*, 273-276. New York: ACM

Forlines, C., Vogel, D., and Balakrishnan, R. 2006b. HybridPointing: fluid switching between absolute and relative pointing with a direct input device. *Proceedings of UIST 2006 – the ACM Symposium on User Interface Software and Technology*, 211-220. New York: ACM

Geiβler, J. 1998. Shuffle, throw or take it! Working Efficiently with an Interactive Wall. *Proceedings of the CHI'98 Conference on Human Factors in Computing Systems,* 265-266. New York: ACM

Graham, E. and MacKenzie, C.L. Pointing on a computer display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI'95*, ACM, New Your, NY, 314-315.

Greenberg, S., Boyle, M. and LaBerge, J. 1999. PDAs and Shared Public Displays: Making Personal Information Public, and Public Information Personal. *Personal Technologies, Vol.3, No.1,* 54-64.

Greenberg, S. and Buxton, B. 2008. Usability evaluation considered harmful (some of the time). In *Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy, April 05 - 10, 2008). CHI '08. ACM, New York, NY, 111-120.

Grossman, T. and Balakrishnan, R. 2005. The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA, April 02 - 07, 2005). CHI '05. ACM, New York, NY, 281-290.

Grudin, J. 2001. Partitioning digital worlds: focal and peripheral awareness in multiple monitor use. *Proceedings of the CHI'01 Conference on Human Factors in Computing Systems,* 458-465. New York: ACM.

Gustafson, S., Baudisch, P., Gutwin, C., and Irani, P. 2008. Wedge: clutter-free visualization of off-screen locations. *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing system*s, ACM (2008), 787-796.

Gutwin, C. 1997. Workspace Awareness in Real-Time Distributed Groupware. *Ph.D. Dissertation*, *Department of Computer Science, University of Calgary*.

Gutwin, C. and Greenberg, S. 1996. Workspace awareness for groupware. In *Conference Companion on Human Factors in Computing Systems: Common Ground* (Vancouver, British Columbia, Canada, April 13 - 18, 1996). M. J. Tauber, Ed. CHI '96. ACM, New York, NY, 208-209.

Gutwin, C. and Greenberg, S. 1998. Design for individuals, design for groups: tradeoffs between power and workspace awareness. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work* (Seattle, Washington, United States, November 14 - 18, 1998). CSCW '98. ACM, New York, NY, 207-216.

Gutwin, C. and Greenberg, S. 2002. A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Comput. Supported Coop. Work* 11, 3 (Nov. 2002), 411-446.

Ha, V., Inkpen, K., Wallace, J., and Ziola, R. 2006. Swordfish: user tailored workspaces in multi-display environments. *Extended Abstracts of the CHI'06 Conference on Human Factors in Computing Systems,* 1487-1492. New York: ACM.

Ha, V., Wallace, J., Ziola, R., and Inkpen, K. 2006b. My MDE: configuring virtual workspaces in multi-display environments. *Extended Abstracts of the CHI'06 Conference on Human Factors in Computing Systems,* 1481-1486. New York: ACM.

Hancock, M. and Carpendale, S. 2007. Supporting Multiple Off-Axis Viewpoints at a Tabletop Display. In *Proceedings of the Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'07)* 171-178, *IEEE*.

Hancock, M., Nacenta, M.A., Gutwin, C., Carpendale, S. 2009. The Effects of Changing Projection Geometry on the Interpretation of 3D Orientation on Tabletops. In *Proceedings of Interactive Tabletops and Surfaces conference 2009. ITS'09*, To appear.

Hart, S. G., & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload*, 139–183. Amsterdam: Elsevier.

Harter, A., Hopper, A., Steggles, P., Ward, A., and Webster, P. 1999. The anatomy of a context-aware application. In *Proceedings of the 5th Annual ACM/IEEE international Conference on Mobile Computing and Networking (MobiCom '99)*, 59-68. New York: ACM.

Hascoët, M. 2003. Throwing models for large displays. In *Proceedings of HCI´2003*, British HCI Group, 73-77.

Haué, J-B., Dillenbourg, P. 2009. Do Fewer Laptops Make a Better Team? In *Interactive Artifacts and Furniture Supporting Collaborative Work and Learning*. Springer. 1-24.

Hazas, M., Kray, C., Gellersen, H., Agbota, H., Kortuem, G., and Krohn, A. 2005. A relative positioning system for co-located mobile devices. *Proceedings of the MobiSys '05 - 3rd international Conference on Mobile Systems, Applications, and Services*, 177-190. New York: ACM.

Heath, C., and Luff, P., 1992. Crisis management and multimedia technology in London Underground Line Control Rooms. In *Journal of Computer Supported Cooperative Work (CSCW)* vol 1, num 1-2 (March, 1992).

Hill, A., Johnson, A. 2008. "Withindows: A Framework for Transitional Desktop and Immersive User Interfaces," *3D User Interfaces, 2008. 3DUI 2008. IEEE Symposium on* , vol., no., pp.3-10, 8-9 March 2008

Hinckley, K. 2003. Synchronous Gestures for Multiple Persons and Computers. In *Proceedings of the ACM symposium on User Interface Software and Technology*. UIST'03, 149-158.

Hinckley, K. 2006. Input Technologies and Techniques. In *Handbook of Human-Computer Interaction*, ed. by Andrew Sears and Julie A. Jacko. Lawrence Erlbaum & Associates.

Hinckley, K., Ramos, G., Guimbretiere, F., Baudisch, P., and Smith, M. (2004) Stitching: pen gestures that span multiple displays. *Proceedings of AVI '04 Conference on Advanced Visual interfaces*, 23-31. New York: ACM

Hinckley, K., Zhao, S., Sarin, R., Baudisch, P., Cutrell, E., Shilman, M., and Tan, D. 2007. InkSeine: *In Situ* search for active note taking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA, April 28 - May 03, 2007). CHI '07. ACM, New York, NY, 251-260.

Holmer T., Streitz, N.A., Geisler, J. 1998. Roomware for cooperative buildings: Integrated design of architectural spaces and information spaces. *Proceedings of CoBuild'98, First International Workshop on Cooperative Building*, 4-21. Heidelberg: Springer.

Holmquist, L.E., Mattern, F., Schiele, B., Alahuhta, P., Beigl, M., and Gellersen,H.-W. 2001. Smart-Its Friends: A Technique for Users to Easily Establish Connections between Smart Artefacts. *Proceedings of Ubicomp 2001*, 116-121. London: Springer-Verlag.

Hornecker, E., Marshall, P., Dalton, N.S., and Rogers, Y. 2008. Collaboration and interference: awareness with mice or touch input. Proceedings of the ACM 2008 conference on Computer supported cooperative work, New York, ACM. 167-176.

Hutchings, D. R. and Stasko, J. 2004. Revisiting display space management: understanding current practice to inform next-generation design. *Proceedings of the 2004 Conference on Graphics interface*, 127-134. New York: ACM.

Hutchings, D. R., Smith, G., Meyers, B., Czerwinski, M., and Robertson, G. (2004) Display space usage and window management operation comparisons between single monitor and multiple monitor users. *Proceedings of the Working Conference on Advanced Visual interfaces (AVI '04),* 32-39. New York: ACM.

Inkpen, K.M., Hawkey, K., Kellar, M., Mandryk, R.L., Parker, J.K., Reilly, D., Scott, S.D. and Whalen, T. 2005. Exploring Display Factors that Influence Co-Located Collaboration: Angle, Size, Number, and User Arrangement. *Proceedings of HCI International 2005*.

Intersense, Inc. 2009. Precision Motion Tracking Solutions. http://www.intersense.com/ Last accessed 10th May 2009.

Ishii, H. and Ullmer, B. 1997. Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, United States, March 22 - 27, 1997). S. Pemberton, Ed. CHI '97. ACM, New York, NY, 234-241.

Izadi, S., Brignull, H., Rodden, T., Rogers, Y., and Underwood, M. 2003. Dynamo: a public interactive surface supporting the cooperative sharing and exchange of media. *Proceedings of UIST 2003 – the ACM Symposium on User Interface Software and Technology*, 159-168. New York: ACM

IT University of Copenhagen 2009. ITU Gaze Tracker. http://www.gazegroup.org/component/content/article/6-news/26-itu-gaze-tracker-released. Last accessed May 10th, 2009.

Jacob, R. J., Sibert, L. E., McFarlane, D. C., and Mullen, M. P. (1994). Integrality and separability of input devices. *ACM Trans. Comput.-Hum. Interact*. 1, 1 (Mar. 1994), 3-26

Ji, Y., Biaz, S., Pandey, S., and Agrawal, P. 2006. ARIADNE: a dynamic indoor signal map construction and localization system. In Proceedings of the 4th international Conference on Mobile Systems, Applications and Services (MobiSys '06). 151-164. New York: ACM.

Jones, W. P. and Dumais, S. T. 1986. The spatial metaphor for user interfaces: experimental tests of reference by location versus name. *ACM Trans. Inf. Syst.* 4, 1 (Jan. 1986), 42-63.

Johanson, B., Hutchins, G., Winograd, T., and Stone, M. 2002. PointRight: experience with flexible input redirection in interactive workspaces. *Proceedings of UIST 2002 – the ACM Symposium on User Interface Software and Technology*, 227-234. New York: ACM.

Johanson, B., Ponnekanti, S., Sengupta, C., Fox, A. 2001. Multibrowsing: Moving Web Content across Multiple Displays. *Proceedings of Ubicomp 2001*, 346-353. London: Springer-Verlag.

Khan, A., Fitzmaurice, G., Almeida, D., Burtnyk, N., and Kurtenbach, G. 2004. A remote control interface for large displays. In *Proceedings of the 17th Annual ACM Symposium on User interface Software and Technology* (Santa Fe, NM, USA, October 24 - 27, 2004). UIST '04. ACM, New York, NY, 127-136.

Kohtake, N., Ohsawa, R., Yonezawa, T., Matsukura, Y., Iwai, M., Takashio, K. and Tokuda, H. 2005. u-Texture: Self-organizable Universal Panels for Creating Smart Surroundings. *Proceedings of the 7th International Conference on Ubiquitous Computing (UbiComp 2005),* 19–36. Heidelberg: Springer

Kohtake, N., Rekimoto, J., and Anzai, Y. 1999. InfoStick: An Interaction Device for Inter-Appliance Computing. *Proceedings of the 1st international Symposium on Handheld and Ubiquitous Computing,* 246-258. London: Springer-Verlag.

Kohtake, N., Rekimoto, J., and Anzai, Y. 2001. InfoPoint: A Device that Provides a Uniform User Interface to Allow Appliances to Work Together over a Network. *Personal Ubiquitous Computing 5, 4 (Jan. 2001),* 264-274

Kornblum, S. (1992) Dimensional overlap and dimensional relevance in stimulus-response and stimulus-stimulus compatibility. *Advances in psychology, vol. 87,* 743-777.

Kornblum, S., Hasbroucq, T., and Osman, A. 1990. Dimensional Overlap: Cognitive Basis for Stimulus-Response Compatibility-A model and Taxonomy. *Psychological Review 1990, Vol 97, No 2,* 253-270.

Kortuem, G., Kray, C., and Gellersen, H. 2005. Sensing and visualizing spatial relations of mobile devices. In *Proceedings of the 18th Annual ACM Symposium on User interface Software and Technology* (Seattle, WA, USA, October 23 - 26, 2005). UIST '05. ACM, New York, NY, 93-102.

Lanir, J., Booth, K. S., and Tang, A., 2008. MultiPresenter: A Presentation System for (Very) Large Display Surfaces. In *Proceedings of the 16th ACM international Conference on Multimedia* (MULTIMEDIA 2008), ACM Press, 519-528.

Leigh, J., Johnson, A., Park, K., Nayak, A., Singh, R. and Chowdry, V. 2002. Amplified Collaboration Environments. *Proceedings of VizGrid Symposium*.

Li, Y., Hinckley, K., Guan, Z., and Landay, J. A. 2005. Experimental analysis of mode switching techniques in pen-based user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA, April 02 - 07, 2005). CHI '05. ACM, New York, NY, 461-470.

MacIntyre, B., Mynatt, E. D., Voida, S., Hansen, K. M., Tullio, J., and Corso, G. M. 2001. Support for multitasking and background awareness using interactive peripheral displays. *Proceedings of UIST 2001 – the ACM Symposium on User Interface Software and Technology*, 41-50. New York: ACM.

Mackenzie, I. S. 1991 *Fitts' Law as a Performance Model in Human-Computer Interaction*. Doctoral Thesis. UMI Order Number: UMI Order No. GAXNN-65985., University of Toronto.

MacKenzie, I. S. and Buxton, W. 1992. Extending Fitts' law to two-dimensional tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Monterey, California, United States, May 03 - 07, 1992). P. Bauersfeld, J. Bennett, and G. Lynch, Eds. CHI '92. ACM, New York, NY, 219-226.

MacKenzie, I. S., Kauppinen, T., and Silfverberg, M. 2001. Accuracy measures for evaluating computer pointing devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, United States). CHI '01. ACM, New York, NY, 9-16.

Mackinlay, J. D. and Heer, J. 2004. Wideband displays: mitigating multiple monitor seams. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria, April 24 - 29, 2004). CHI '04. ACM, New York, NY, 1521-1524.

Mandryk, R. L. 2004. Objectively evaluating entertainment technology. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria, April 24 - 29, 2004). CHI '04. ACM, New York, NY, 1057-1058.

Mandryk, R.L. and Gutwin, C. Perceptibility and utility of sticky targets. 2008. *Proceedings of graphics interface 2008*, Canadian Information Processing Society (2008), 65-72.

Mandryk, R. L., Rodgers, M. E., and Inkpen, K. M. 2005. Sticky widgets: pseudo-haptic widget enhancements for multi-monitor displays. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (Portland, OR, USA, April 02 - 07, 2005). CHI '05. ACM, New York, NY, 1621-1624.

Manesis, T. and Avouris, N. 2005. Survey of position location techniques in mobile systems. In *Proceedings of the 7th international Conference on Human Computer interaction with Mobile Devices & Services (MobileHCI '05)*, vol. 111, 291-294. New York: ACM.

Mantei, M. 1988. Capturing the capture concepts: a case study in the design of computer-supported meeting environments. In *Proceedings of the 1988 ACM Conference on*

Computer-Supported Cooperative Work (Portland, Oregon, United States, September 26 - 28, 1988). CSCW '88. ACM, New York, NY, 257-270.

Marshall, P., Hornecker, E., Morris, R., Dalton, N.S., and Rogers, Y. 2008. When the Fingers do the Talking: A Study of Group Participation with Varying Constraints to a Tabletop Interface. 3rd IEEE International Workshop on Horizontal Interactive Human Computer Systems, TABLETOP 2008, 33-40.

Massó, J. P., Vanderdonckt, J., and López, P. G. 2006. Direct manipulation of user interfaces for migration. *Proceedings of the 11th international Conference on intelligent User interfaces (IUI '06),* 140-147. New York: ACM.

McGrath, J. E., & Hollingshead, A. B. 1993. *Groups interacting with technology: ideas, evidence, issues and an agenda*. Sage Publications, Inc. Thousand Oaks, CA, USA.

McGrath, J. E. 1995. Methodology matters: doing research in the behavioral and social sciences. In *Human-Computer interaction: Toward the Year 2000*, R. M. Baecker, J. Grudin, W. A. Buxton, and S. Greenberg, Eds. Morgan Kaufmann Publishers, San Francisco, CA, 152-169.

Meyer, D., Abrams, R., Kornblum, S., Wright, C., & Smith, J. Optimality. 1988. Human motor performance: Ideal control of rapid aiming movements. In *Psychological Review*, 95, (1988), 340-370.

Microsoft Corporation. 2008. *Multimonitor Support in Microsoft Windows Vista.* http://www.microsoft.com/whdc/device/display/multimonVista.mspx Last accessed June 7th 2008.

Microsoft Corporation. 2008b. *How to configure file sharing in Windows XP.* http://support.microsoft.com/kb/304040 Last accessed June 7th 2008.

Microsoft Corporation 2009. *Surface.* http://www.microsoft.com/surface/. Last accessed May 10th 2009.

Microsoft Corporation. 2008c. *HOW TO: Use Windows Messenger Instant Messaging.* http://support.microsoft.com/kb/307887 Last accessed June 8th 2008.

Miller, R. C. and Myers, B. A. (1999) Synchronizing clipboards of multiple computers. *Proceedings of UIST 1999 – the ACM Symposium on User Interface Software and Technology*, 65-66. New York: ACM

Morris, M. R., Huang, A., Paepcke, A., and Winograd, T. 2006. Cooperative gestures: multi-user gestural interactions for co-located groupware. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada, April 22 - 27, 2006). R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, and G. Olson, Eds. CHI '06. ACM, New York, NY, 1201-1210.

Moyle, M., Cockburn, A. 2002. Analysing Mouse and Pen Flick Gestures. *Proceedings of the SIGCHI-NZ Symposium On Computer-Human Interaction,* 19-24

Myers, B. A., Bhatnagar, R., Nichols, J., Peck, C. H., Kong, D., Miller, R., and Long, A. C. 2002. Interacting at a distance: measuring the performance of laser pointers and other devices. *Proceedings of the SIGCHI Conference on Human Factors in Computing System*. CHI '02. ACM Press, New York, NY, 33-40.

Myers, B.A., Peck, C.H., Nichols, J., Kong, D., and Miller, R. 2001. Interacting at a Distance Using Semantic Snarfing. *Proc. UbiComp 2001*, 305-314.

Myers, B. A., Stiel, H., and Gargiulo, R. 1998. Collaboration using multiple PDAs connected to a PC. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work* (Seattle, Washington, United States, November 14 - 18, 1998). CSCW '98. ACM, New York, NY, 285-294.

Nacenta, M.A., Aliakseyeu, D., Subramanian, S., and Gutwin, C. 2005. A comparison of techniques for Multi-Display Reaching. *Proceedings of the CHI'05 Conference on Human Factors in Computing Systems,* 371-380. New York: ACM

Nacenta, M.A., Pinelle, D., Stuckel, D., and Gutwin, C. 2007. The effects of interaction technique on coordination in tabletop groupware. In *Proceedings of Graphics interface 2007* (Montreal, Canada, May 28 - 30, 2007). GI '07, vol. 234. ACM, New York, NY, 191-198.

Nacenta, M.A., Sakurai, S., Yamaguchi, T., Miki, Y., Itoh, Y., Kitamura, Y., Subramanian, S. and Gutwin, C. 2007b. E-conic: a Perspective-Aware Interface for Multi-Display Environments. *Proceedings UIST 2007 – the ACM Symposium on User Interface Software and Technology*, 279-288. New York, ACM.

Nacenta, M.A., Gutwin, C., Aliakseyeu, D., and Subramanian, S. 2009. There and Back Again: Cross-Display Object Movement in Multi-Display Environments. In journal of *Human-Computer Interaction,* 24:1, 170-229.

Ni, T., Schmidt, G.S., Staadt, O.G., Livingston, M.A., Ball, R., and May, R. 2006. A Survey of Large High-Resolution Display Technologies, Techniques, and Applications. Virtual Reality Conference, IEEE, IEEE Computer Society, 223-236.

Norman, D. (2002) *The design of Everyday Things*. Basic Books.

Nunamaker, J. F., Dennis, A. R., Valacich, J. S., Vogel, D., and George, J. F. 1991. Electronic meeting systems. *Commun. ACM* 34, 7 (Jul. 1991), 40-61.

Oh, J.Y. & Stuerzlinger, W. 2002. Laser pointers as collaborative pointing devices. *Proceedings of Graphics Interfaces (GI'02),* 141-149.

Olsen, D. R. and Nielsen, T. 2001. Laser pointer interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, United States). CHI '01. ACM, New York, NY, 17-22.

Parker, J. K., Mandryk, R. L, Inkpen, K. M. 2005. TractorBeam: Seamless Integration of Local and Remote Pointing for Tabletop Displays. *Proc. Graphics Interface* (2005), 33-40.

Parker, K., Mandryk, R., Nunes, M., Inkpen, K. 2005. TractorBeam Selection Aids: Improving Target Acquisition for Pointing Input on Tabletop Displays. *Proceedings of INTERACT 2005 - the Tenth IFIP TC13 International Conference on Human-Computer Interaction*, 80-93. Heidelberg: Springer.

Patten, J., Ishii, H., Hines, J., and Pangaro, G. 2001. Sensetable: a wireless object tracking platform for tangible user interfaces *Proceedings of the CHI'01 Conference on Human Factors in Computing Systems,* 253-260. New York: ACM.

Pinelle, D., Dyck, J., Gutwin, C., Stach, T. 2006. Cutouts: Multiple Views for Tabletop Groupware., *Technical Report HCI-TR-06-04*,

Pinelle, D., Barjawi, M., Nacenta, M., and Mandryk, R. 2009. An evaluation of coordination techniques for protecting objects and territories in tabletop groupware. In *Proceedings of the 27th international Conference on Human Factors in Computing Systems* (Boston, MA, USA, April 04 - 09, 2009). CHI '09. ACM, New York, NY, 2129-2138.

Pinelle, D., Nacenta, M., Gutwin, C., and Stach, T. 2008. The effects of co-present embodiments on awareness and collaboration in tabletop groupware. In *Proceedings of Graphics interface 2008* (Windsor, Ontario, Canada, May 28 - 30, 2008). GI, vol. 322. Canadian Information Processing Society, Toronto, Ont., Canada, 1-8.

Plamondon, R. A kinematic theory of rapid human movements. *Biological Cybernetics* 72, 4 (1995), 309-320.

Po, B.A., Fisher, D., and Booth, S. 2005. Comparing cursor orientations for mouse, pointer, and pen interaction. *Proceedings of the CHI'05 Conference on Human Factors in Computing Systems*, 291-300. New York: ACM.

Polhemus, 2009. Motion Tracking. http://www.polhemus.com/?page=Motion_Liberty_Latus. Last accessed May 10th, 2009.

Ponnekanti, S. R., Johanson, B., Kiciman, E., and Fox, A. 2003. Portability, Extensibility and Robustness in iROS. In *Proceedings of the First IEEE international Conference on Pervasive Computing and Communications (PERCOM)*. IEEE Computer Society, Washington DC, 11.

Prante, T., Streitz, N., and Tandler, P. 2004. Roomware: Computers Disappear and Interaction Evolves. *Computer* 37, 12 (Dec. 2004), 47-54.

Proctor, R. W. and Vu, K. L. 2003. Human information processing: an overview for human-computer interaction. In J. A. Jacko and A. Sears (Eds.), *the Human-Computer interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, Human Factors and Ergonomics* (pp. 35-51). Mahwah, NJ: Lawrence Erlbaum Associates.

Proctor, R.W. & Reeve, T.G. 1990. *Stimulus-response compatibility: an integrated perspective*. New York: North-Holland.

Ramos, G., Boulos, M., and Balakrishnan, R. 2004. Pressure widgets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vienna, Austria, April 24 - 29, 2004). CHI '04. ACM, New York, NY, 487-494.

Randall, J., Amft, O., Bohn, J., Burri, M. 2007. LuxTrace - indoor positioning using building illumination. *Personal and Ubiquitous Computing Journal*, 11(6), 417-428. Springer.

Realtime Soft. UltraMon. 2009. http://www.realtimesoft.com/ultramon/ Last accessed Jun 20-06-2009.

Reetz, A., Gutwin, C., Stach, T., Nacenta, M., and Subramanian, S. 2006. Superflick: a natural and efficient technique for long-distance object placement on digital tables. *Proceedings of the GI'06 conference on Graphics interface*, 163-170. New York: ACM.

Rekimoto, J. 1997. Pick-and-Drop A Direct Manipulation Technique for Multiple Computer Environments. *Proceedings of UIST 1997 – the ACM Symposium on User Interface Software and Technology*, 31-39. New York, ACM.

Rekimoto, J., Ayatsuka, Y., and Kohno, M. 2003. SyncTap: An interaction technique for mobile networking. In *Proceedings of the Mobile Human Computer Interaction Conference 2003*, 104-115. Berlin: Springer.

Rekimoto, J., Ayatsuka, Y., Kohno, M., and Oba, H. 2003b. Proximal Interactions: A Direct Manipulation Technique for Wireless Networking. *Proceedings of INTERACT 2003 - the Ninth IFIP TC13 International Conference on Human-Computer Interaction,* 511-518. Amsterdam: IOS Press

Rekimoto, J. and Saitoh, M. 1999. Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments, *Proceedings of the CHI'99 Conference on Human Factors in Computing Systems*, 378-385. New York: ACM

Ringel, M. (2003). When one isn't enough: an analysis of virtual desktop usage strategies and their implications for design. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA, April 05 - 10, 2003). CHI '03. ACM, New York, NY, 762-763.

Robertson, S., Wharton, C., Ashworth, C. & Franzke, M. 1996. Dual Device User Interface Design: PDAs and Interactive Television. *Proceedings of the CHI'96 Conference on Human Factors in Computing Systems,* 79-86. New York: ACM.

Rogers, Y., Lim, Y-K., Hazlewood, W. R. & Marshall, P. 2009. Equal Opportunities: Do Shareable Interfaces Promote More Group Participation Than Single User Displays?. Human-Computer Interaction, 24 (1), 79-116.

Rogers,Y. & Lindley,S. 2004. Collaborating around vertical and horizontal large interactive displays: which way is best? *Interacting with Computers, Volume 16, Issue 6,* 1133-1152.

Rogers, Y. and Rodden, T. 2003. Configuring spaces and surfaces to support collaborative interactions. In O'Hara, K., Perry, M., Churchill, E. and Russell, D. (Eds.) *Public and Situated Displays*, Kluwer Publishers.

Román, M., Hess, C., Cerqueira, R., Ranganathan, A., Campbell, R. H., and Nahrstedt, K. 2002. A Middleware Infrastructure for Active Spaces. *IEEE Pervasive Computing* 1, 4 (Oct. 2002), 74-83.

Schmidt, R., Zelaznik, H., Hawkins, B., Frank, J., & Quinn, J. 1979. Motor output variability: A theory for the accuracy of rapid motor acts. *Psych. Review*, 86, (1979), 415-451.Shiffrin, R. M., & Dumais, S. T. 1981. The development of automatism. In J. R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, NJ: Lawrence Erlbaum.

Scott, S. D., Sheelagh, M., Carpendale, T., and Inkpen, K. M. 2004. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA, November 06 - 10, 2004). CSCW '04. ACM, New York, NY, 294-303.

Sears, A. and Shneiderman, B. 1991. High precision touchscreens: design strategies and comparisons with a mouse. *Int. J. Man-Mach. Stud.* 34, 4 (Apr. 1991), 593-613.

Shell, J. S., Selker, T., and Vertegaal, R. 2003. Interacting with groups of computers. *Communications of the ACM 46, 3,* 40-46.

Shoemaker, G., Tang, A., and Booth, K. S. 2007. Shadow reaching: a new perspective on interaction for large displays. In *Proceedings of the 20th Annual ACM Symposium on User interface Software and Technology* (Newport, Rhode Island, USA, October 07 - 10, 2007). UIST '07. ACM, New York, NY, 53-56.

Siio, I. 1995. InfoBinder: a pointing device for a virtual destop system. *Proceedings of the Sixth International Conference on Human-Computer Interaction 1995 v.III,* 261-264. Elsevier Science

Slay, H., Thomas, B., and Vernik, R. 2003. An interaction model for universal interaction and control in multi display environments. *Proceedings of the 1st international Symposium on information and Communication Technologies,* 220-225. New York: ACM

Slay, H., Thomas, B., Vernik, R., and Piekarski, W. 2004. A Rapidly Adaptive Collaborative Ubiquitous Computing Environment to Allow Passive Detection of Marked Objects. *Proceedings of APCHI'04, Asia Pacific CHI,* 420-430. Heidelberg: Springer.

Soukoreff, R.W. and MacKenzie, I.S. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *International Journal of Human-Computer Studies* 61, 6 (2004), 751-789.

Spratt, M. 2003. An Overview of Positioning by Diffusion. *Wireless Networks*, Kluwer Academic Publishers, November. 2003, 9 (6), 565 – 574.

Stefik, M., Bobrow, D. G., Lanning, S., Tatar, D., and Foster, G. 1986. WYSIWIS revisited: early experiences with multi-user interfaces. *Proceedings of the 1986 ACM Conference on Computer-Supported Cooperative Work (CSCW '86)*, 276-290. New York: ACM

Stoakley, R., Conway, M., Pausch, R. (1995) Virtual reality on a WIM: interactive worlds in miniature. *Proceedings of the CHI'95 Conference on Human Factors in Computing Systems,* 265-272. New York: ACM.

Streitz, N. A., Geißler, J., Holmer, T., Konomi, S., Müller-Tomfelde, C., Reischl, W., Rexroth, P., Seitz, P., and Steinmetz, R. 1999. i-LAND: an interactive landscape for creativity and innovation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: the CHI Is the Limit* (Pittsburgh, Pennsylvania, United States, May 15 - 20, 1999). CHI '99. ACM, New York, NY, 120-127.

Su, R.E. and Bailey, B.P. 2005. Put Them Where? Towards Guidelines for Positioning Large Displays in Interactive Workspaces. In *Proceedings of the IFIP conference on Human-Computer Interaction - INTERACT 2005*. LNCS 3585. Springer. Pp. 337-349.

Swaminathan, K. and Sato, S. 1997. Interaction design for large displays. *interactions* 4, 1 (Jan. 1997), 15-24.

Swindells, C. 2002. *Use that there! Pointing to determine device identity.* M.Sc. Thesis, Simon Fraser University, Vancouver, Canada.

Swindells, C., Inkpen, K., Dill, J., Tory, M. 2002. That one there! Pointing to establish device identity. *Proceedings of UIST 2002 – the ACM Symposium on User Interface Software and Technology*, 151-160. New York: ACM.

Synergy, 2009. http://synergy2.sourceforge.net/ Last accessed 20-06-2009.

Tan, D. & Czerwinski, M. 2003. Effects of Visual Separation and Physical Discontinuities when Distributing Information across Multiple Displays. *Proceedings of OZCHI 2003 Conference for the Computer-Human Interaction Special Interest Group of the Ergonomics Society of Australia*, 184-191.

Tan, D. S., Meyers, B., and Czerwinski, M. 2004. WinCuts: manipulating arbitrary window regions for more effective use of screen space. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria, April 24 - 29, 2004). CHI '04. ACM, New York, NY, 1525-1528.

Tang, A., Tory, M., Po, B., Neumann, P., and Carpendale, S. 2006. Collaborative coupling over tabletop displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada, April 22 - 27, 2006). R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, and G. Olson, Eds. CHI '06. ACM, New York, NY, 1181-1190.

Tang, J.C. 1991. Findings from observational studies of collaborative work. In *International Journal of Man-Machine Studies,* 34, pp. 143-160.

Tandler, P. 2002. Architecture of beach: The software infrastructure for roomware environments. In *CSCW 2000: Workshop on Shared Environments to Support Face-to-Face Collaboration*, Philadelphia, PA, December 2002.

Tandler, P., Prante, T., Müller-Tomfelde, C., Streitz, N., and Steinmetz, R. (2001) Connectables: dynamic coupling of displays for the flexible creation of shared workspaces. *Proceedings of UIST 2001 – the ACM Symposium on User Interface Software and Technology*, 11-20. New York: ACM.

Tobii Technology. 2009. Tobii eye-tracking hardware. http://www.tobii.com/market_research_usability/products_services/eye_tracking_hardware.aspx. Last accessed May 10th, 2009.

Ullmer, B., Ishii, H., and Glas, D. 1998. media-Blocks: Physical containers, transports, and controls for online media. *Proceedings of SIGGRAPH'98 Conference on Computer graphics and interactive techniques*, 379-386. New York: ACM

Vicon Motion Systems, 2009. Motion Tracking applications. http://www.vicon.com/applications/simulation.html. Last accessed May 10th, 2009.

Vildjiounaite, E., Malm, E., Kaartinen, J., and Alahuhta, P. (2002). Location Estimation Indoors by Means of Small Computing Power Devices, Accelerometers, Magnetic Sensors, and Map Knowledge. In *Proceedings of the First international Conference on Pervasive Computing*, 211-224. London: Springer-Verlag.

Voida, S., Podlaseck, M., Kjeldsen, R., and Pinhanez, C. 2005. A study on the manipulation of 2D objects in a projector/camera-based augmented reality environment. *Proceedings of* *the CHI'01 Conference on Human Factors in Computing Systems,* 611-620. New York: ACM.

Wacom corporation. *Wacom Penabled technology* http://www.wacom-components.com/english/technology/penabled.html. Last Accessed June 14th 2008.

Wallace, J. R., Mandryk, R. L., and Inkpen, K. M. 2008. Comparing content and input redirection in MDEs. In *Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work* (San Diego, CA, USA, November 08 - 12, 2008). CSCW '08. ACM, New York, NY, 157-166.

Wigdor, D., Jiang, H., Forlines, C., Borkin, M., and Shen, C. WeSpace: the design development and deployment of a walk-up and share multi-surface visual collaboration system. *Proceedings of the 27th international conference on Human factors in computing systems (CHI'09)* , ACM (2009), 1237-1246.

Wigdor, D., Shen, C., Forlines, C., and Balakrishnan, R. 2006. Table-centric interactive spaces for real-time collaboration. *Proceedings of the Working Conference on Advanced Visual interfaces (AVI '06),* 103-107. New York: ACM.

Wigdor, D., Shen, C., Forlines, C., and Balakrishnan, R. 2007. Perception of elementary graphical elements in tabletop and multi-surface environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA, April 28 - May 03, 2007). CHI '07. ACM, New York, NY, 473-482.

Wikipedia. *Multi-monitor*. http://en.wikipedia.org/wiki/Multi-monitor Last accessed June 7th 2008.

Wilson, A. and Sarin, R. 2007. BlueTable: Connecting Wireless Mobile Devices on Interactive Surfaces Using Vision-Based Handshaking. *Proceedings of Graphics Interface 2007,* 119-125. New York: ACM.

Woodworth, R.S. 1899. The accuracy of voluntary movements, *Psych. Review Monograph*, Supp., 3, 1-114.

Wu, M., Balakrishnan, R. 2003. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. *Proceedings of UIST 2003 – the ACM Symposium on User Interface Software and Technology*, 193-202. New York: ACM.

Zajonc, R. (1965) Social Facilitation. Science, 149 (1965), 269.

Zellweger, P. T., Mackinlay, J. D., Good, L., Stefik, M., and Baudisch, P. 2003. City lights: contextual views in minimal space. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA, April 05 - 10, 2003). CHI '03. ACM, New York, NY, 838-839.

Zhai, S., Morimoto, C., Ihde, S. 1999. Manual and Gaze Input Cascaded (MAGIC) Pointing. *Proceedings of the CHI'99 Conference on Human Factors in Computing Systems,* 246-253. New York: ACM

APPENDIX A: EXPERIMENT CONSENT FORMS, DEMOGRAPHIC DATA FORMS AND POST-STUDY QUESTIONNAIRES

The experiments and the handling of participants of all the experiments described in this dissertation were carried out according to the University of Saskatchewan's guidelines for experiments with human subjects, under approved Ethics applications (investigator, Prof. Carl Gutwin, and investigator Regan Mandryk).

Contents

UNIVERSITY OF SASKATCHEWAN

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF SASKATCHEWAN

CONSENT FORM

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

Researcher(s): *Sriram Subramanian*, Department of Computer Science (966-4888)
Carl Gutwin, Department of Computer Science (966-8646)
Miguel Nacenta, Department of Computer Science (966-6593)

Purpose and Procedure:

This study is concerned with understanding the role of spatial information in virtual representations.

The goal of the research is to determine if faithful representation of the physical space in the virtual space makes interaction easier or faster.
The session will require 50 to 70 minutes, during which you will be asked to carry out several target acquisition tasks.
At the end of the session, you will be given more information about the purpose and goals of the study, and there will be time for you to ask questions about the research.

There is no known risk to you associated with this study.

The data collected from this study will be used in articles for publication in journals and conference proceedings.

As one way of thanking you for your time, we will be pleased to make available to you a summary of the results of this study once they have been compiled (probably in about four weeks). This summary will outline the research and discuss our findings and recommendations. If you would like to receive a copy of this summary, please write down your email address here.

Contact email address____

Confidentiality: All of the information we collect from you (data logged by the computer, observations made by the experimenters and your questionnaire responses) will be stored so that your name is not associated with it (using an arbitrary participant number). Any write-ups of the data will not include any information that can be linked directly to you. Please do not put your name or other identifying information on the questionnaire. The research materials will be stored with complete security throughout the entire investigation. Do you have any questions about this aspect of the study?

Right to Withdraw: You are free to withdraw from the study at any time without penalty and without losing any advertised benefits, including the 10$. Withdrawal from the study will not affect your academic status or your access to services at the university. If you withdraw, your data will be deleted from the study and destroyed. In addition, you are free to not answer specific items or questions on questionnaires.

Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact: Sriram Subramanian (966-4888)

Questions: If you have any questions concerning the study, please feel free to ask at any point; you are also free to contact the researchers at the numbers provided above if you have questions at a later time. This study has been approved on ethical grounds by the University of Saskatchewan Behavioural Sciences Research Ethics Board on November 2003. Any questions regarding your rights as a participant may be addressed to that committee through the Office of Research Services (966-2084). Out of town participants may call collect.

Consent to Participate: *I have read and understood the description provided above; I have been provided with an opportunity to ask questions and my questions have been answered satisfactorily. I consent to participate in the study described above, understanding that I may withdraw this consent at any time. A copy of this consent form has been given to me for my records.*

______________________________ ____________________

(Signature of Participant) (Date)

____

(Signature of Researcher)

Participant demographic sheet (Exp 1 & 2)

User: **A B C D** Group: _____ Day: ____ September 2006

Age: _____

Gender: Female / Male

Do you have any kind of visual impairment (e.g., color-blindness, non-corrected myopia etc..?)
Yes / No __________

Have you used a tablet-PC before? Yes / No
For how long? _______

Have you used a PDA before Yes / No
For how long? _______

How many hours a week do you use a computer? _______

Which applications do you use most often?

□ Word processors or text editors (Word, WordPerfect, MS Works, NotePad)
□ E-mail
□ Internet Browsers
□ Graphic Editing Software (Photoshop, Paint, Paint Shop Pro, Illustrator…)
□ Video Editing Software
□ Programming environments (Eclipse, Visual Studio…)
□ Games
□ Other ____

How many hours a week do you play computer games? ________

If you play games at all, which kind?

□ First-person shooters
□ Strategy
□ Tetris-like
□ Simulation (car racing, flight simulation)
□ Board Games (Computer Risk, Monopoly etc)
□ Other __________

Workload assessment form (part 1-anamorphic first) (Exp 2)

User: **A B C D**
Group:
Day:
We need you to tell us about the demands of the task you have just finished in each one of the situations: alone, matching, and non-matching.
Please, rate the following aspects of the task **when you were doing it by yourself,** at the very beginning of the experiment.

Mental Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Physical Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Temporal Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Own Performance
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Effort
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Frustration
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Turn the page to see the rest of the questionnaire.

Please, rate the following aspects of the task **when the positions of the names in the computer DID NOT match the people in the real world.**

Mental Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Physical Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Temporal Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Own Performance
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Effort
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Frustration
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Please, rate the following aspects of the task **when the positions of the names in the computer matched the people in the real world.**

Mental Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Physical Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Temporal Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Own Performance
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Effort
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Frustration
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Finally, if you had to perform all the trials again and could choose in which situation to repeat them, what would you choose:

- ☐ With the positions of the names **NOT** matching the actual positions of the people
- ☐ With the positions of the names matching the actual positions of the people
- ☐ I don't care

Why? (please ask for more paper if you need more space)

Workload assessment form (part 2-anamorphic first) (Exp 1)

User: **A B C D** Group: Day:

We need you to tell us about the demands of the task in each of the two situations: matching, and non-matching.

Please, rate the following aspects of the task **when the positions of the names in the computer DID NOT match the people in the real world.**

Mental Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Physical Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Temporal Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Own Performance
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Effort
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Frustration
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Please, rate the following aspects of the task **when the positions of the names in the computer matched the people in the real world.**

Mental Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Physical Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Temporal Demand
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Own Performance
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Effort
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Frustration
Low 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 **High**

Please, turn the page for the rest of the questionnaire.

Finally, if you had to perform all the trials again and could choose in which situation to repeat them, what would you choose:

- ☐ With the positions of the names **NOT** matching the actual positions of the people
- ☐ With the positions of the names matching the actual positions of the people
- ☐ I don't care

Why?

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF SASKATCHEWAN
CONSENT FORM

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

Researcher: Sriram Subramanian, Department of Computer Science (966-4888)
Student: Miguel Nacenta, Department of Computer Science (966-2327)

Purpose and Procedure:

This study is concerned with finding next generation input devices for smart meeting rooms.

The goal of the research is to determine which pen-based input technique is best for transferring objects between displays.
The session will require about 60 minutes, during which you will be asked to carry out several trials of a task consisting on transferring a digital object from one point in one display to another point from another display.
At the end of the session, you will be given more information about the purpose and goals of the study, and there will be time for you to ask questions about the research.

There is no known risk to you associated with this study.

The data collected from this study will be used in articles for publication in journals and conference proceedings.

As one way of thanking you for your time, we will be pleased to make available to you a summary of the results of this study once they have been compiled (probably in about four weeks). This summary will outline the research and discuss our findings and recommendations. If you would like to receive a copy of this summary, please write down your email address here.

Contact email address: ____

Confidentiality: All of the information we collect from you (data logged by the computer, observations made by the experimenters, your questionnaire responses and video or audio recording) will be stored so that your name is not associated with it (using an arbitrary participant number). If audio or video recording is used for data collection, the recording will be such that your personal identity is not compromised. Any write-ups of the data will not include any information that can be linked directly to you. Please do not put your name or other identifying information on the questionnaire. The research materials will be stored with complete security throughout the entire investigation. Do you have any questions about this aspect of the study?

Right to Withdraw: You are free to withdraw from the study at any time without penalty and without losing any advertised benefits. Withdrawal from the study will not affect your academic status or your access to services at the university. If you withdraw, your data will be deleted from the study and destroyed. If audio or video recording is used for data collection, you have the right to switch off the audio or video recorder at any time during the study. In addition, you are free to not answer specific items or questions on questionnaires.

Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact: Sriram Subramanian (966-4888)

Questions: If you have any questions concerning the study, please feel free to ask at any point; you are also free to contact the researchers at the numbers provided above if you have questions at a later time. This study has been approved on ethical grounds by the University of Saskatchewan Behavioural Sciences Research Ethics Board on 20th Nov. 2003 for 5 years. Any questions regarding your rights as a participant may be addressed to that committee through the Office of Research Services (966-2084). Out of town participants may call collect.

Consent to Participate: *I have read and understood the description provided above; I have been provided with an opportunity to ask questions and my questions have been answered satisfactorily. I consent to participate in the study described above, understanding that I may withdraw this consent at any time. A copy of this consent form has been given to me for my records.*

______________________________ ____________________

(Signature of Participant) (Date)

____

(Signature of Researcher)

Questionnaire about Interaction Techniques subject # 1 (Exp 3, part 1)

Put in order of preference the techniques that you have used, i.e., number 1 is the one you liked the most, number 2 the second, and so on (they are in the order they were presented to you, if you don't remember any of them, ask the experimenter).

__ Sync Gestures
__ Radar
__ Slingshot
__ Pantograph
__ Pick & Drop
__ Press & Flick

Put in order of speed the techniques that you have used, i.e., the one that you think is the fastest will be number one, and so on (again ask if you don't remember the name of a technique).

__ Sync Gestures
__ Radar
__ Slingshot
__ Pantograph
__ Pick & Drop
__ Press & Flick

Put the techniques in order from most to least tiring.

__ Sync Gestures
__ Radar
__ Slingshot
__ Pantograph
__ Pick & Drop
__ Press & Flick

Put the techniques in order from giving you most control to least control.

__ Sync Gestures
__ Radar
__ Slingshot
__ Pantograph
__ Pick & Drop
__ Press & Flick

Now please tell us what are the most obvious advantages and disadvantages from the techniques (you don't have to fill in all of them):

Sync Gestures

Advantages

Disadvantages

Radar

Advantages

Disadvantages

Slingshot

Advantages

Disadvantages

Pantograph

Advantages

Disadvantages

Pick & Drop

Advantages

Disadvantages

Press & Flick

Advantages

Disadvantages

Questionnaire about Interaction Techniques (Exp 3, part 2)

Put in order of preference the techniques that you have used, i.e., number 1 is the one you liked the most, number 2 the second, and so on (they are in the order they where presented to you, if you don't remember any of them, ask the experimenter).

__ Pantograph
__ Slingshot
__ Radar
__ Press & Flick

Put in order of speed the techniques that you have used, i.e., the one that you think is the fastest will be number one, and so on (again ask if you don't remember the name of a technique).

__ Pantograph
__ Slingshot
__ Radar
__ Press & Flick

Put the techniques in order from most to least tiring.

__ Pantograph
__ Slingshot
__ Radar
__ Press & Flick

Put the techniques in order from giving you most control to least control.

__ Pantograph
__ Slingshot
__ Radar
__ Press & Flick

Now please tell us what are the most obvious advantages and disadvantages from the techniques (you don’t have to fill in all of them):

__ Pantograph
Advantages

Disadvantages

__ Slingshot
Advantages

Disadvantages

__ Radar
Advantages

Disadvantages

__ Press & Flick
Advantages

Disadvantages

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF SASKATCHEWAN
CONSENT FORM

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

Researcher(s): *Sriram Subramanian*, Department of Computer Science (966-4888)
Miguel Nacenta, Department of Computer Science

Purpose and Procedure:

This study is concerned with understanding the role of perspective in multi-display reaching.

The goal of the research is to determine if controlling the pointer with knowledge of the users' perspective improves reaching times in multi-display environments.
The session will require 60 to 75 minutes, during which you will be asked to carry out several target acquisition tasks.
At the end of the session, you will be given more information about the purpose and goals of the study, and there will be time for you to ask questions about the research.

There is no known risk to you associated with this study.

The data collected from this study will be used in articles for publication in journals and conference proceedings.

As one way of thanking you for your time, we will be pleased to make available to you a summary of the results of this study once they have been compiled (probably in about four weeks). This summary will outline the research and discuss our findings and recommendations. If you would like to receive a copy of this summary, please write down your email address here.

Contact email address____

Confidentiality: All of the information we collect from you (data logged by the computer, observations made by the experimenters and your questionnaire responses) will be stored so that your name is not associated with it (using an arbitrary participant number). Any write-ups of the data will not include any information that can be linked directly to you. Please do not put your name or other identifying information on the questionnaire. The research materials will be stored with complete security throughout the entire investigation. Do you have any questions about this aspect of the study?

Right to Withdraw: You are free to withdraw from the study at any time without penalty and without losing any advertised benefits, including the 10$. Withdrawal from the study will not affect your academic status or your access to services at the university. If you withdraw, your data will be deleted from the study and destroyed. If audio or video recording is used for data collection, you have the right to switch off the audio or video recorder at any time during the study. In addition, you are free to not answer specific items or questions on questionnaires.

Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact: Sriram Subramanian (966-4888)

Questions: If you have any questions concerning the study, please feel free to ask at any point; you are also free to contact the researchers at the numbers provided above if you have questions at a later time. This study has been approved on ethical grounds by the University of Saskatchewan Behavioural Sciences Research Ethics Board on November 2003. Any questions regarding your rights as a participant may be addressed to that committee through the Office of Research Services (966-2084). Out of town participants may call collect.

Consent to Participate: *I have read and understood the description provided above; I have been provided with an opportunity to ask questions and my questions have been answered satisfactorily. I consent to participate in the study described above, understanding that I may withdraw this consent at any time. A copy of this consent form has been given to me for my records.*

__________________________________ ____________________

(Signature of Participant) (Date)

____

(Signature of Researcher)

Subject Information

Subject Id:	____________
Sex:	M/F
Age:	____________
Handedness:	R/L
Major:	____________
Use of computers (hours/week)	____________
Programs:	____
Use of games (hours/week)	____________
Games:	____
Use of graphical app. (hours/week)	____________
Graph. Apps	____
Ball sports practice(days/month)	____________
Ball sports:	____________
Shooting, darts?	____________
Beam Performance:	____________

Workload Assessment

Stitched Displays

Mental Demand

Low High

Physical Demand

Low High

Temporal Demand

Low High

Effort

Low High

Performance

Good Poor

Frustration

Low High

Perspective Cursor

Mental Demand

Low High

Physical Demand

Low High

Temporal Demand

Low High

Effort

Low High

Performance

Good Poor

Frustration

Low High

Laser Pointer

Mental Demand

Low High

Physical Demand

Low High

Temporal Demand

Low High

Effort

Low High

Performance

Good Poor

Frustration

Low High

Post-study questionnaire

Which technique do you think was faster? And second?

Which technique do you think was more accurate? And second?

Which technique do you think was the best? And second?

General Comments

UNIVERSITY OF SASKATCHEWAN

DEPARTMENT OF COMPUTER SCIENCE

UNIVERSITY OF SASKATCHEWAN
CONSENT FORM

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

Researcher(s): *Carl Gutwin,* Department of Computer Science (966-2327)
Miguel A. Nacenta, Department of Computer Science

Purpose and Procedure:

This study is concerned with interaction in multi-display environments. The goal of the research is to improve the methods that multi-display systems offer to users.

The session will require 40 to 60 minutes, during which you will be asked to carry out repeated targeting tasks. At the end of the session, you will be given more information about the purpose and goals of the study, and there will be time for you to ask questions about the research. As compensation for your time you will receive $10.

There is no known risk to you associated with this study.

The data collected from this study will be used in articles for publication in journals and conference proceedings. As one way of thanking you for your time, we will be pleased to make available to you a summary of the results of this study once they have been compiled (probably in about four weeks). This summary will outline the research and discuss our findings and recommendations. If you would like to receive a copy of this summary, please write down your email address here.

Contact email address____

Confidentiality: All of the information we collect from you (data logged by the computer, observations made by the experimenters and your questionnaire responses) will be stored so that your name is not associated with it (using an arbitrary participant number). Any write-ups of the data will not include any information that can be linked directly to you. Please do not put your name or other identifying information on the questionnaire. The research materials will be stored with securely throughout the entire investigation. Do you have any questions about this aspect of the study?

Right to Withdraw: You are free to withdraw from the study at any time without penalty and without losing any advertised benefits, including the 10$. Withdrawal from the study will not affect your academic status or your access to services at the university. If you withdraw, your data will be deleted from the study and destroyed. If audio or video recording is used for data collection, you have the right to switch off the audio or video recorder at any time during the study. In addition, you are free to not answer specific items or questions on questionnaires.

Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact: Miguel Nacenta (966-6593)

Questions: If you have any questions concerning the study, please feel free to ask at any point; you are also free to contact the researchers at the numbers provided above if you have questions at a later time. This study has been approved on ethical grounds by the University of Saskatchewan Behavioural Sciences Research Ethics Board on November 2003. Any questions regarding your rights as a participant may be addressed to that committee through the Office of Research Services (966-2084). Out of town participants may call collect.

Consent to Participate: *I have read and understood the description provided above; I have been provided with an opportunity to ask questions and my questions have been answered satisfactorily. I consent to participate in the study described above, understanding that I may withdraw this consent at any time. A copy of this consent form has been given to me for my records.*

________________________________ ____________________

(Signature of Participant) (Date)

____

(Signature of Researcher)

Participant Demographic Information

Participant ID: __________
Age: _____
Sex: □ Female □ Male
Major or main professional activity: _____________
How many hours a week do you spend using a computer (average)? ________
Do you play games regularly? If so, how many hours a week (average)? _____
What kinds of games do you play and how many hours a week? (e.g., first person shooters (3), strategy (15), Sims (4))

____

Which of the games that you play show the whole game world in the screen all the time? (e.g., pakman, tetris, chess)

____

Do you use regularly a multi-monitor setting? How many hours per week? ________________________

Have you ever used a multi-monitor setting? If so, for how long (hours total)?_____________________

Please, draw the physical setup of the multi-monitor environments that you have used the most (Draw only monitors table, keyboard and mice – or any other input and output devices).

Subjective Evaluation Form

Participant Id: _______

Remember there were three multi-monitor connection options:
Jump: The cursor jumps from one monitor to the next. The cursor is always visible at least in one monitor.
Travel: The cursor travels through the space between monitors. The cursor is not visible when it is in between monitors.
Travel with halos: The cursor travels through the space between monitors. The position of the cursor is indicated by red circles

Please indicate your preference of connection options for each of the distances. Please, mark only one box in each category (that is, only mark one technique as best, one as intermediate and one as worst).

SHORT DISTANCE:

Jump	□ Best	□ Intermediate	□ Worst
Travel	□ Best	□ Intermediate	□ Worst
Travel with halos	□ Best	□ Intermediate	□ Worst

Why?

MEDIUM DISTANCE:

Jump	□ Best	□ Intermediate	□ Worst
Travel	□ Best	□ Intermediate	□ Worst
Travel with halos	□ Best	□ Intermediate	□ Worst

Why?

LONG DISTANCE:

Jump	□ Best	□ Intermediate	□ Worst
Travel	□ Best	□ Intermediate	□ Worst
Travel with halos	□ Best	□ Intermediate	□ Worst

Why?

Please indicate which connection option allowed you to be fastest.

SHORT DISTANCE:

Jump	□ Best	□ Intermediate	□ Worst
Travel	□ Best	□ Intermediate	□ Worst
Travel with halos	□ Best	□ Intermediate	□ Worst

Why?

MEDIUM DISTANCE:

Jump	□ Best	□ Intermediate	□ Worst
Travel	□ Best	□ Intermediate	□ Worst
Travel with halos	□ Best	□ Intermediate	□ Worst

Why?

LONG DISTANCE:

Jump	□ Best	□ Intermediate	□ Worst
Travel	□ Best	□ Intermediate	□ Worst
Travel with halos	□ Best	□ Intermediate	□ Worst

Why?

Please indicate which connection option allowed you to be most accurate.

SHORT DISTANCE:

Jump	□ Best	□ Intermediate	□ Worst
Travel	□ Best	□ Intermediate	□ Worst
Travel with halos	□ Best	□ Intermediate	□ Worst

Why?

MEDIUM DISTANCE:

Jump	□ Best	□ Intermediate	□ Worst
Travel	□ Best	□ Intermediate	□ Worst
Travel with halos	□ Best	□ Intermediate	□ Worst

Why?

LONG DISTANCE:

Jump	□ Best	□ Intermediate	□ Worst
Travel	□ Best	□ Intermediate	□ Worst
Travel with halos	□ Best	□ Intermediate	□ Worst

Why?

Finally, did you have concerns about losing the mouse with the **travel** techniques? How did it affect you?

APPENDIX B: CLASSIFICATION OF TECHNIQUES ACCORDING TO A CDOM IT TAXONOMY

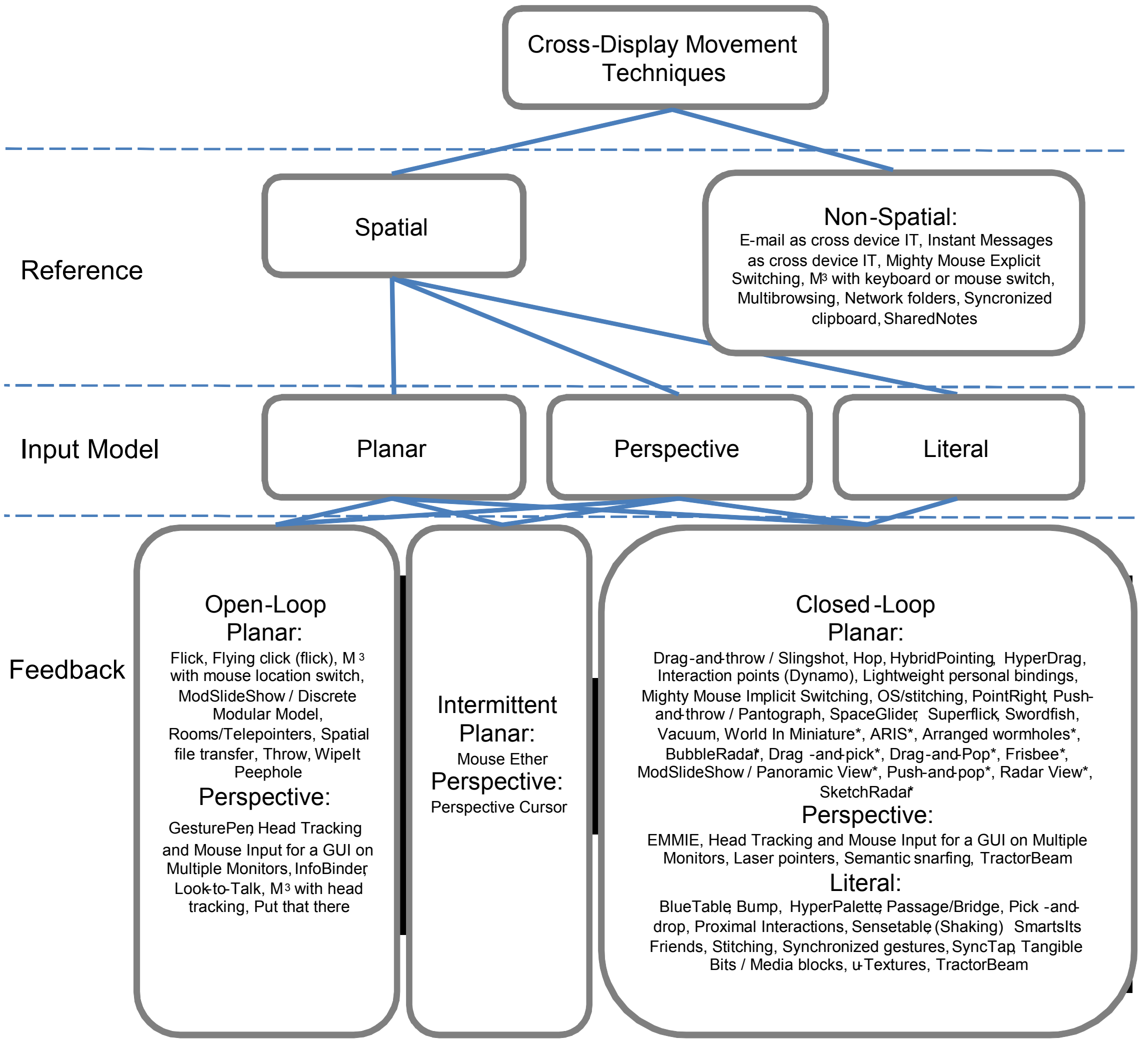

Figure 69. A classification of existing interaction techniques according to the taxonomy derived from the categories of Chapter 3, 4, 5 and 6. For the purpose of this classification, world-in-miniature techniques (those marked with "*") are considered closed-loop because they afford absolute control of the objects in the miniature. However, these techniques only provide feedback for the full-size objects in the environment when the object is in display space. For other users, or depending on the requirement of the task, these techniques should be considered intermittent.

The following table contains an alphabetical list of interaction techniques. In some cases, the techniques are not named, and instead the name of the system appears. The last column references to the source where the technique is define or, by default, to a source where it is compared or studied. Note that many of the techniques presented here were not designed or evaluated for multi-display scenarios; for completeness, the list contains include all techniques that could be used for CDOM, regardless of whether they were motivated in that way by their original authors.

Abbreviations:

Referential domain (2nd column): SP – spatial
NS – non-spatial

Input model (3rd column): PL – planar
L – literal
PR – perspective
NA – Not applicable

Feedback (4th column): O – open-loop
C – closed-loop
C* - closed-loop (world-in-miniature - see below)
IN – intermittent

Note: for the purpose of this classification, world-in-miniature techniques are considered closed-loop because they afford absolute control of the objects in the miniature. However, these techniques only provide feedback for the full-size objects in the environment when the object is in display space. For other users, depending on the requirement of the task, these techniques should be considered intermittent (IN).

NAME	**Referential Domain**	**Input Model**	**Feedback**	**Reference**
ARIS' application relocation	**SP**	**PL**	**C***	(Biehl & Bailey, 2006)
Arranged wormholes	**SP**	**PL**	**C***	(Wu & Balakrishnan, 2003)
BlueTable	**SP**	**L**	**C**	(Wilson & Sarin, 2007)
BubbleRadar	**SP**	**PL**	**C***	(Aliakseyeu et al., 2006)
Bump	**SP**	**L**	**C**	(Hinckley, 2003)
Drag-and-Pick	**SP**	**PL**	**C***	(Baudisch et al., 2003)
Drag-and-Pop	**SP**	**PL**	**C***	(Baudisch et al., 2003)
Drag-and-Throw / Slingshot	**SP**	**PL**	**C**	(Hascoët, 2003)
E-mail as cross device IT	**NS**	**NA**	**O**	-
EMMIE	**SP**	**PR**	**C**	(Butz et al., 1999)
Flick	**SP**	**PL**	**O**	(Moyle & Cockburn, 2002)
Flying click (flick)	**SP**	**PL**	**O**	(Dulberg et al., 2003)
Frisbee	**SP**	**PL**	**C***	(Khan et al., 2003)
GesturePen	**SP**	**PR**	**O**	(Swindells et al., 2002)

Head Tracking and Mouse Input for a GUI on Multiple Monitors	**SP**	**PR**	**O**	(Ashdown et al., 2005)
HybridPointing	**SP**	**PL**	**C**	(Forlines et al., 2006b)
HyperDrag	**SP**	**PL**	**C**	(Rekimoto & Saitoh, 1999)
HyperPalette	**SP**	**L**	**C**	(Ayatsuka et al., 2000)
InfoBinder	**SP**	**PR**	**O**	(Siio, 1995)
InfoStick/InfoPoint	**SP**	**PR**	**C**	(Kohtake et al., 1999; Kohtake et al., 2001)
Instant Messages as cross device IT	**NS**	**NA**	**O**	-
Interaction points (Dynamo)	**SP**	**PL**	**C**	(Izadi et al., 2003)
Laser pointers	**SP**	**PR**	**C**	(Olsen & Nielsen, 2001)
Lightweight personal bindings	**SP**	**PL**	**C**	(Ha et al., 2006b)
Mighty Mouse Explicit Switching	**NS**	**NA**	**O**	(Booth et al., 2002)
Mighty Mouse Implicit Switching	**SP**	**PL**	**C**	(Booth et al., 2002)
M^3 with head tracking	**SP**	**PR**	**O**	(Benko & Feiner, 2005)
M^3 with keyboard Switch	**NS**	**NA**	**O**	(Benko & Feiner, 2005)
M^3 with mouse location switch	**SP**	**PL**	**O**	(Benko & Feiner, 2005)
M^3 withMouse Button Switch	**NS**	**NA**	**O**	(Benko & Feiner, 2005)
ModSlideShow / Discrete Modular Model	**SP**	**PL**	**O**	(Chiu et al., 2003)
ModSlideShow / Panoramic View	**SP**	**PL**	**C***	(Chiu et al., 2003)
Mouse Ether	**SP**	**PL**	**IN**	(Baudisch et al., 2004)
Multibrowsing	**NS**	**NA**	**O**	(Johanson et al., 2001)
Network folders	**NS**	**NA**	**O**	-
OS/stitching	**SP**	**PL**	**C**	http://www.microsoft.com/windowsxp/using/setup/learnmore/northrup_multimon.mspx
Passage/Bridge	**SP**	**L**	**C**	(Streitz et al., 1999)
Perspective Cursor	**SP**	**PR**	**IN**	(Nacenta et al., 2006)
Pick-and-Drop	**SP**	**L**	**C**	(Rekimoto, 1997)
PointRight	**SP**	**PL**	**C**	(Johanson et al., 2002)
Proximal Interactions	**SP**	**L**	**C**	(Rekimoto et al., 2003b)
Push-and-pop	**SP**	**PL**	**C***	(Collomb et al., 2005)
Push-and-throw / Pantograph	**SP**	**PL**	**C**	(Hascoët, 2003)
Put-that-There	**SP**	**PR**	**O**	(Bolt, 1980)
Radar View	**SP**	**PL**	**C***	(Swaminathan and Sato, 1997)
Rooms/Telepointers	**SP**	**PL**	**O**	(Stefik et al., 1986)
Semantic snarfing	**SP**	**PR**	**C**	(Myers et al., 2001)
Sensetable	**SP**	**L**	**C**	(Patten et al., 2001)
(Shaking) Smarts-Its Friends	**SP**	**L**	**C**	(Holmquist et al., 2001)
Syncronized clipboard	**NS**	**NA**	**O**	(Miller & Myers, 1999)
SharedNotes	**NS**	**NA**	**O**	(Greenberg et al., 1999)
SketchRadar	**SP**	**PL**	**C***	(Aliakseyeu & Martens, 2006)
SpaceGlider	**SP**	**PL**	**C**	(Leigh et al., 2002)
Spatial file transfer	**SP**	**PL**	**O**	(Hazas et al., 2005)
Stitching	**SP**	**L**	**C**	(Hinckley et al., 2004)
Superflick	**SP**	**PL**	**C**	(Reetz et al., 2006)

Swordfish	**SP**	**PL**	**C**	(Ha et al., 2006-1)
Synchronized gestures	**SP**	**L**	**C**	(Nacenta et al., 2005)
SyncTap	**SP**	**L**	**C**	(Rekimoto et al., 2003)
Tangible Bits / Media blocks	**SP**	**L**	**C**	(Ullmer et al., 1998)
Throw	**SP**	**PL**	**O**	(Geiβler, 1998)
TractorBeam	**SP**	**PR/L**	**C**	(Parker et al., 2005)
u-Textures	**SP**	**L**	**C**	(Kohtake et al., 2005)
Vacuum	**SP**	**PL**	**C**	(Bezerianos & Balakrishnan, 2005)
WipeIt Peephole	**SP**	**PL**	**O**	(Butz & Krüger, 2006)
World in Miniature	**SP**	**PL**	**C***	(Stoakley et al., 1995)